WORKING PAPER

CLIMATE CHANGE CONCERNS AND THE PERFORMANCE OF GREEN VERSUS BROWN STOCKS

David Ardia
Keven Bluteau
Kris Boudt
Koen Inghelbrecht

March 2021
2021/1011

GHENT UNIVERSITY

Department of Economics

Faculty of Economics and Business Administration
Campus Tweekerken, St.-Pietersplein 5, 9000 Ghent - BELGIUM

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

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

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

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

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

[d] *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: February 26, 2021

Abstract

We empirically test the prediction of Pastor, Stambaugh, and Taylor (2020) 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 Media Climate Change Concerns index using news about climate change published by major U.S. newspapers. We find that when concerns about climate change increase unexpectedly, green firms' stock prices increase, while brown firms' decrease. Further, using topic modeling, we conclude that climate change concerns affect returns both through investors updating their expectations about firms' future cash flows and through changes in investors' preferences for sustainability.

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 Laurent Barras and participants at the Quantact seminar 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@hec.ca` (Keven Bluteau), `kris.boudt@UGent.be` (Kris Boudt), `koen.inghelbrecht@UGent.be` (Koen Inghelbrecht)

1. Introduction

Many consider climate change to be 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 prices of financial assets (Fama and French, 2007). Anecdotal evidence suggests that preference shifts have caused a 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.*, Climate Action Plan).[1]

Pastor, Stambaugh, and Taylor (2020) 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 customer 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. Thus, an increase in investors' preferences for green assets because of increasing concerns about climate change increases (decreases) the discount rate of brown

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

(green) firms leading to a decrease (increase) in stock prices. This paper empirically tests the prediction of Pastor, Stambaugh, and Taylor (2020) that green firms outperform brown firms when concerns about climate change increase unexpectedly.

The challenge in testing the above is that unexpected changes in concerns about climate change is latent and must be proxied. Engle et al. (2020) use news media articles to build two monthly indices to proxy for climate change risk. The first index captures the *attention* about climate change in the Wall Street Journal. The second index relies on the Crimson Hexagon proprietary sentiment measure to capture the *negative attention* about climate change.[2] Similarly, we use media news data but aim at capturing *concerns* about climate change. To do so, we propose a novel "concerns score" measuring the level of negativity as well as the level of risk and uncertainty discussed in each article. We rely on news from eight major and highly circultated U.S. newspapers, including the Wall Street Journal. Following Baker, Bloom, and Davis (2016), we combine the daily concerns' scores, considering heterogeneity across sources, into a daily Media Climate Change Concerns index (MCCC). Finally, we obtain a proxy of unexpected changes in climate change concerns using the prediction error of a first-order autoregressive model calibrated on the MCCC index, which we refer to as unexpected media climate change concerns (UMC). Overall, compared to Engle et al. (2020), our index pinpoints concerns in articles by combining attention and informational content (*i.e.*, uncertainty and sentiment) about climate change portrayed in the news media, is computed at a higher frequency, and is freely available.[3] With our methodology, it is also straightforward to build topical indices to isolate specific themes about climate change concerns such as natural disasters, as shown later in our paper.

Our empirical study focuses on S&P 500 firms from January 2010 to June 2018. To quantify a firm's greenness, we rely on 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, *i.e.*, the number of tonnes of CO2-

[2]The attention index is available at `https://sites.google.com/view/stefanogiglio/` but the negative attention index is not freely available.

[3]The MCCC index is available at `https://sentometrics-research.com/`.

equivalent GHG emissions necessary for a firm to generate $1 million in revenue. Firms below the 25th percentile for this variable on a given day are defined as green firms, and firms above the 75th percentile are defined as brown firms.

We first analyze the contemporaneous relationship between UMC and the daily return of a green-minus-brown (GMB) portfolio that is long in green firms and short in brown firms. We find a significant positive relationship, suggesting that green stocks can outperform brown stocks when there are unexpected increases in climate change concerns. When looking at the green (brown) portfolio returns individually, we find a positive (negative) and significant relationship with UMC. This relationship 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 they reward green firms. Moreover, we also find that neutral firms (firms that are classified as neither green nor brown) have a positive relationship with UMC, albeit to a lesser degree than green firms. These findings are consistent with the observation of Bolton and Kacperczyk (2020) that institutional investors tend to screen firms on direct emissions intensity in a few salient industries, and then reallocate capital to other firms, which can be neutral or green.

Next, we use panel regressions to estimate the exposure of individual firms' stock returns to UMC, conditional on their emissions intensity. Our results are in line with our previous findings: The lower (higher) the emissions intensity, the more positive (negative) the exposure to unexpected increases in climate change concerns. In related work, Ilhan, Sautner, and Vilkov (2020) show that the variation in GHG emissions intensity is to a large extend explained by the industry. Hence, we test whether the exposure to UMC is still driven by the GHG emissions intensity of the firm when removing the industry effect and find it is the case. As an additional robustness check, we test whether firms that do not disclose their GHG emissions are affected by unexpected changes in climate change concerns. We find that there is no significant difference between non-disclosing and disclosing firms.

In the final analysis, we investigate whether the stock price reaction to UMC arises from expectations about firms' cash flows or changes to investor preferences, as predicted in the model of Pastor, Stambaugh, and Taylor (2020). We extract general themes discussed in the climate change news article data and build topical MCCC indices. We then evaluate which channel is more likely to be affected for each theme, conditional on its effect on green vs. brown stock performance. Our analysis identifies eight themes (*i.e.*, clusters of topics) related to climate change, of which five have a significant relationship with green vs. brown firms' stock performance: (i) Financial and Regulation, (ii) Agreements and Summits, (iii) Societal Impact, (iv) Research and (v) Disasters. Among these, we posit that the Financial and Regulation theme primarily affects the cash flow channel. Conversely, the Research and Disaster themes are likely to affect the investor tastes channel. Finally, the Agreement and Summit and Societal Impact themes may affect both channels. Overall, our results suggest that the effect arises from changes to both cash flow expectations and investor tastes.

By empirically verifying the predictions of Pastor, Stambaugh, and Taylor (2020) using our new daily MCCC index, we complement several recent studies in the literature that focus on understanding the impact of climate change on financial markets. 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. 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. Ramelli et al. (2018) 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. Bertolotti et al. (2019) analyze the impact of extreme weather events on U.S. electric utilities' stock prices. Bolton and Kacperczyk (2020) study whether carbon emissions affect the cross-section of the U.S. stock market. Görgen et al. (2020) develop and study a carbon risk factor using a long-short portfolio based on a carbon emissions-related measure.

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 relationship 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 Pastor, Stambaugh, and Taylor (2020), 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.[4] One could derive unexpected changes from this survey, in particular unexpected changes in the answer to the question about how worried participants are about global warming or climate change. However, this survey (as well as others) is conducted very infrequently, limiting the measure's usefulness. Instead, we proxy ΔCC_t on a daily basis using news media data.

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 for this proxying.

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.*, see Schoenfeld, Meier, and Griffin, 1979; Slovic,

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

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 the people's attitudes towards a subject. Based on this theory, the level of concerns 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.[5] 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

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

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: A risk lexicon to determine the level of discussion about (future) risk-events and a sentiment lexicon to assess the increase in (the perception of) risk. These lexicons are retrieved from the LIWC2015 software (Pennebaker et al., 2015).[6] 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.[7]

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 $N_{n,t,s}$ in a news article n published on day t by source s, the article's concerns score is defined as:

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

The first ratio of the product, $(\frac{RW_{n,t,s}}{N_{n,t,s}})$, 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' length. The second ratio, $\big(\frac{NW_{n,t,s}-PW_{n,t,s}}{NW_{n,t,s}+PW_{n,t,s}} + 1\big)/2$, 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-

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

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

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 for 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 (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 subsection 2.1. Note that when no news is published about climate change (*i.e.*,$N_{t,s} = 0$), the concerns score in (3) is 0, which is equivalent to a 100% positive sentiment term in (2). As such, our approach assumes that no news is good news.[8]

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 ($1 \leq \tau_1 < \tau_2 \leq T$):

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

[8] 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 decreases over time and increases in the presence of information about the damage of climate change. Thus, they make a similar assumption that no news is good news.

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 news that agrees with their views, limiting the reach of alternative information to these groups (for example, see Flaxman, Goel, and Rao, 2016). Another argument comes from the concept of "opinion inertia," which arises, for instance, from the confirmation bias (for example, see 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.[9]

2.4. Unexpected changes in the Media Climate Change Concerns index

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, $\Delta CC_t - \mathbb{E}[\Delta CC_t|I_{t-1}]$. Because the media tends to publish unexpected information, it is reasonable to use $MCCC_t$ as a 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). To account for "expected" news, we estimate a first-order autoregressive model on $MCCC_t$ and interpret the prediction error as the unexpected changes in climate change concerns.[10] We refer to the prediction error as UMC_t in the remainder of the paper. More details are provided in subsection 3.2.

2.5. Comparison with existing methodologies

Our index construction is related to the methodologies presented in Engle et al. (2020), which proposes two ways to capture climate risk from the news. A first approach relies on Wall Street Journal news articles, and a lexicon referred to as the "Climate Change Vocabulary" derived from authoritative texts about climate change. The method extracts a similarity feature between each news article in the corpus and the Climate Change Vocabulary. The higher the similarity measure, the more likely it is that an article discusses climate change. This similarity feature is then aggregated on a monthly basis to obtain a climate change risk index. The rationale with this measure is that media attention on climate change can proxy for the risk level, as the media will only report about sufficiently important climate change news. In our framework, this assumption is valid under

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

[10] We also consider the prediction error of an ARMA(p,q) model with lags selected with the Bayesian information criterion (Schwarz et al., 1978) as a robustness check, and obtain qualitatively similar results.

the condition that most of the content in these news articles expresses concerns. We also consider media attention as a component to capture changes in climate change concerns. Their second approach relies on the natural language proprietary algorithms of Crimson Analytics to compute online news articles' negative sentiments about climate change. We also use textual sentiments, but compute it with a lexicon. We leverage the negative sentiments in our article-level concerns score, using it as a weight for the article-level risk score.

Our index construction combines both the level of attention and the informational content about climate change portrayed in the news media. Moreover, by including a textual risk measure in our article-level concerns score, we posit that we more accurately proxy for unexpected changes in climate change concerns than using only a single dimension of the available information (*e.g.*, attention, sentiment or risk).[11]

3. Data

Our study relies on climate change news articles published by multiple sources, and 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 from January 1, 2003, to June 30, 2018.[12] 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.[13] We consider sources with a daily circulation of more than 500,000 newspapers: (i) Wall Street

[11]We find that the correlation between an index that only uses average concerns ($\sum_{n=1}^{N_{t,s}} \frac{concerns_{n,t,s}}{N_{t,s}}$) and an index that only uses attention ($N_{t,s}$) is 77%, indicating that media attention is also related to concerns.

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

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

Journal, (ii) New York Times, (iii) Washington Post, (iv) Los Angeles Times, (v) Chicago Tribune, (vi) USA Today, (vii) New York Daily News, and (viii) New York Post.[14]

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.[15] We filter out short news articles with fewer than 200 words, as lexicon-based methods are typically noisy for short texts.

In Table 1, we report the number of climate change articles, the total number of news articles, and the percentage of climate change news articles published by the sources in our sample. The source that publishes the most about climate change in terms of the number of articles is the Wall Street Journal, with 3,776 articles. The New York Times publishes the most relative to its total number of articles (0.25%). The Chicago Tribune, New York Daily News and New York Post publish 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 (509 vs. 249), USA Today publishes more about climate change in relative terms than the Chicago Tribune (0.17% vs. 0.05%). This heterogeneity highlights that standardization by sources before aggregation is necessary, as each newspaper appears to have a different focus.

[Insert Table 1 about here.]

For the 20 articles with the highest concerns scores in our corpus, Table 2 reports (i) the publication date, (ii) the concerns score, (iii) the level of risk, (iv) the level of negativity, (v) the first 50 characters of the article's headline and (vi) the source. From the headline, we see that the most concerning articles appear to be legitimately concerning.[16]

[14] We do not examine the Houston Chronicle or Arizona Republic, despite having a daily circulation of more than 500,000 newspapers, as they are not included in the databases used in this study.

[15] 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.

[16] We do not report the least concerning articles, as many of the articles have a concerns score of zero, thus lacking any apparent relevance to climate change.

[Insert Table 2 about here.]

To get a better overview of climate change topics discussed in our corpus, 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, which infers latent correlated topics among a collection of texts.[17] 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 attribution $\theta_{k,n,t,s}$ 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 $\omega_{v,k} \geq 0$. We estimate the model with $K = 40$ topics; more details are provided in Appendix A.

In Table 3, we report the ten words or collocations (*i.e.*, common sequences of two words) with the highest probability for each topic (*i.e.*, the ten largest $\omega_{v,k}$ for each topic k). We also organize the topics into eight clusters that constitute more general themes for ease of interpretation; see Appendix A for details. From these clusters, we see that climate change discussions in the news media is spread across several themes, which we label as: (i) "Financial and Regulation," (ii) "Agreement and Summit," (iii) "Societal Impact," (iv) "Research," (v) "Disaster," (vi) "Environmental Impact," (vii) "Agricultural Impact" and (viii) "Other."

[Insert Table 3 about here.]

To better understand how much attention the media devotes to these topics over time, we compute the number of monthly article equivalents for each 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.

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

In Figure 1, we display the monthly number of article equivalents for each theme from January 2010 to June 2018. The most discussed themes (in decreasing order) are: "Financial and Regulation," "Agreement and Summit," "Societal Impact," "Research," "Disaster," "Environmental" and "Agricultural Impact." We observe significant time variations in the percentage of coverage devoted to each theme. For instance, the "Agreement and Summit" theme tends to have a larger number of article equivalents during months when there are notable conferences on climate change. Similarly, we observe an increase in the "Disaster" theme in 2012 and 2017, which had very destructive wildfire seasons. The time variation of newspapers' coverage of themes implies that each topic captures different dimensions of the climate change discussion.

[Insert Figure 1 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 2, we display the daily evolution of the index from 2003 to 2018. Note that the 2003 to 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.[18]

[Insert Figure 2 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

[18]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.

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 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.[19]

To extract the unexpected component of the MCCC index, we use a first-order autoregressive model. Specifically, at time t, we estimate an AR(1) model with three years of data up to time $t-1$ and use the prediction error for UMC_t.[20]

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. To quantify these damages, we use the greenhouse gas (GHG) emissions disclosed by firms. We retrieve these variables from the Asset4/Refinitiv database. Similar to Ilhan, Sautner, and Vilkov (2020), we focus on S&P 500 firms because surveys of greenhouse gas emissions typically target these firms.[21]

The greenhouse gas emissions variable is separated into three scopes defined by the GHG Protocol Corporate Standard.[22] 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)

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

[20] Results are similar if an expanded window is used instead of a rolling window of three years. Similar results are also obtained using an ARMA(p,q) model with lags chosen with the Bayesian information criterion (BIC).

[21] Our results and conclusions are robust when considering S&P 1500 firms. However, beyond the S&P 500 universe, few firms disclose their greenhouse gas emissions.

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

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.[23] 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, 2020).[24]

In Table 4, we report the percentage of firms in the S&P 500 with available GHG emissions (Panel A) and summary statistics for GHG emissions intensity (Panel B). While our GHG emissions source differs from Ilhan, Sautner, and Vilkov (2020), who use the Carbon Disclosure Project database[25], we see that our coverage of S&P 500 firms is similar, averaging slightly above 50% of the firms in the universe. The average emissions intensity is 682.49 tonnes of CO2-equivalent emissions per $1 million in revenue. The 25th and 75th percentiles are 21.54 and 378.93, respectively. The quartiles, together with the skewness and kurtosis statistics, indicate a distribution of GHG emissions intensity that is highly positively skewed and fat-tailed.

[Insert Table 4 about here.]

Finally, we note that GHG emissions are typically reported with a one-year delay. Similar to Ilhan, Sautner, and Vilkov (2020), we account for this by shifting the GHG emissions intensity variable by 12 months in our analyses.

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

[24] 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.

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

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

We first construct portfolios of green and brown stocks and test whether the green portfolio outperforms the brown portfolio when there are unexpected increases in climate change concerns, both using a conditional mean analysis (Section 4.1) and a multivariate factor analysis (Section 4.2). Next, we analyze the impact of climate change concerns in the cross-section of stock returns (Section 4.3). In particular, we evaluate whether industry-relative GHG emissions intensity matters, and whether firms that do not disclose their emissions are impacted by unexpected changes in climate change concerns.

4.1. Conditional mean analysis

We divide assets into three groups: green, neutral and brown. Green (brown) stocks are firms with a GHG emissions intensity variable in the lowest (highest) quartile of all firms' values on day t. Neutral firms are the remainder of firms that disclose GHG emissions data.[26] We then build, for each day, equal-weighted portfolios for these groups.[27]

Our first analysis focuses on the average return of the green minus brown (GMB) portfolio conditional on the *UMC* variable. In Figure 3, 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 clear positive relationship between the average return and *UMC*. In particular, when *UMC* is above its median, we notice strong increases in the GMB portfolio average return as the thresholds 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 3 about here.]

[26] 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.

[27] While GHG emissions intensity is updated yearly, stocks can enter or exit the S&P500 universe at any day. Also, we note that results are qualitatively similar if market capitalization-weighted portfolios are used instead.

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 common factors used in the financial literature ($\mathbf{f}_t$). We consider the five Fama-French factors (Fama and French, 2015): (i) *MKT*, the excess market return; (ii) *SMB*, the small minus big factor; (iii) *HML*, the high minus low factor; (iv) *RMW*, the robust minus weak factor; and (v) *CMA*, the conservative minus aggressive factor. We also include (vi) *MOM*, the momentum factor of Carhart (1997).[28] This yields the following specification:

$$r_{p,t} = c_p + \beta_p^{UMC} UMC_t + \boldsymbol{\beta}_p \mathbf{f}_t + \varepsilon_{p,t} \,, \tag{7}$$

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 Pastor, Stambaugh, and Taylor (2020) model, we expect that $\beta_{GMB}^{UMC} > 0$, $\beta_G^{UMC} > 0$, and $\beta_B^{UMC} < 0$.

Estimation results are reported in Table 5. First, let us 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 9 basis points. This effect is highly significant, with the t-stat at about 3.3 — above the significance hurdle of 3.0 proposed by Harvey, Liu, and Zhu (2016). The estimated coefficients indicate that the GMB portfolio is positively related to *MKT*, *HML*, *SMB* and *MOM*, and negatively related to *CMA* and *RMW*. Thus, the GMB portfolio emphasizes small firms with lower growth, aggressive investment policies and weak operating profits. The *CMA* coefficient (-0.559) is large compared to the other coefficients. This finding is consistent with green firms investing more and brown firms investing less, which is another implication of the Pastor, Stambaugh, and Taylor (2020) model. This prediction arises from the idea that green firms' capital costs are lower than brown firms'. Thus, more investment opportunities

[28] Factors and risk-free rate data are retrieved from Kenneth French's website at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html`.

for green firms have a positive net present value, resulting in a higher investment level relative to their size than for brown firms.

[Insert Table 5 about here.]

Looking at the green portfolio, we find a positive and highly significant exposure to *UMC*. For the brown portfolio, we find a highly significant negative coefficient. Moreover, we find that the *UMC* coefficient for the brown portfolio is larger in absolute value than for the green portfolio (0.054 vs. 0.037). We also find that neutral firms have a positive relationship with unexpected changes in climate change concerns. However, the coefficient for the neutral portfolio is lower than for the green portfolio (0.022 vs. 0.037). This finding implies that investors' strategies regarding climate change tend toward a screening of brown firms, with reallocation to both green and neutral firms, consistent with Bolton and Kacperczyk (2020).

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

In the previous section, we 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 relationship using stock-level return exposures to *UMC*. Moreover, we test whether the results still hold when we consider variations in GHG emissions intensity within industries, rather than across industries. We also analyze whether firms that do not disclose their GHG emissions are affected by climate change concerns based on their industry, and if this effect differs from firms that disclose their emissions.

4.3.1. General model

We first define $lGHG_{i,t}$ as the cross-sectionally standardized logarithm of the greenhouse gas intensity of firm i available at time t.[29] The standardization is performed by focus-

[29] Similar results are obtained if we use the cross-sectional median for standardization as opposed to the cross-sectional average.

ing on the cross-sectional variation across firms. We then estimate the following panel regression model:

$$r_{i,t} = c + \gamma^{lGHG} lGHG_{i,t} + \left(\gamma^{UMC} + \gamma^{UMC}_{lGHG}\, lGHG_{i,t}\right) UMC_t + \boldsymbol{\beta}_i \mathbf{f}_t + \epsilon_{i,t}\,, \tag{8}$$

where $r_{i,t}$ is the excess stock return of firm i at time t, and $\mathbf{f}_t$ are control factors. We consider one-factor (*MKT*), three-factor (*MKT*, *HML*, *SMB*) and six-factor (*MKT*, *HML*, *SMB*, *RMW*, *CMA*, *MOM*) specifications. Coefficients $\gamma^{\bullet}$ are common to all firms, while $\boldsymbol{\beta}_i$ are firm-specific coefficients.[30]

In specification (8), 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. Given our previous results for the neutral portfolio, we expect a positive value for the common component, γ^{UMC}. Moreover, we expect a significant 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, in line with the prediction by Pastor, Stambaugh, and Taylor (2020).[31]

We also consider an asymmetric specification to test our earlier finding that the firm value exposure to UMC is not a linear function in GHG emissions (*i.e.*, brown firms are more affected by UMC than green firms, and neutral firms are positively affected). To do so, we introduce the variable $A_{i,t}$, which is equal to one when $lGHG_{i,t} > 0$; that is, an indicator variable that is equal to one when the log-GHG emissions intensity of firm i is above the cross-sectional average at time t. As such, this variable captures firms that

[30] In addition, we consider a firm fixed-effects specification where c is replaced by c_i as well as a threshold model in which regression parameters are conditioned on the value of *UMC* being above or below a certain threshold calibrated with the Bayesian information criterion. Our conclusions remained unchanged.

[31] A Fama-MacBeth cross-sectional regression analysis was also performed and provided similar results (see Appendix B).

tends toward the browner side of the green vs. brown spectrum at each point in time. We then estimate the following panel regression model, which nests the previous one:

$$r_{i,t} = c + \left(\gamma^{lGHG} + \delta_A^{lGHG} A_{i,t}\right) lGHG_{i,t} \\ + \left(\gamma^{UMC} + \gamma_{lGHG}^{UMC} lGHG_{i,t} + \delta_{lGHG\text{-}A}^{UMC} lGHG_{i,t} A_{i,t}\right) UMC_t + \boldsymbol{\beta}_i \mathbf{f}_t + \epsilon_{i,t}\,. \quad (9)$$

In this specification, the exposure of firms to unexpected changes in climate change concerns is $\left(\gamma^{UMC} + \gamma_{lGHG}^{UMC} lGHG_{i,t} + \delta_{lGHG\text{-}A}^{UMC} lGHG_{i,t} A_{i,t}\right)$. We expect a negative value for $\delta_{lGHG\text{-}A}^{UMC}$, which would imply that browner firms (*i.e.*, above the cross-sectional average) are more exposed in absolute terms to unexpected changes in climate change concerns than greener firms (*i.e.*, below the cross-sectional average).

Our panel regression models allow us to test the implication of the model by Pastor, Stambaugh, and Taylor (2020); that is, green firms outperform brown firms when there are unexpected increases in climate change concerns. Our asymmetric specification allows us to test the implication of Bolton and Kacperczyk (2020); that is, institutional investors tend to screen for emissions-intense firms, which are clustered in a few salient industries, but do not necessarily prioritize investing in the greenest firms. This observation implies an asymmetry between greener and browner firms' exposures and a positive relationship between neutral firms and *UMC*, as observed in our portfolio analysis.

Panel regression results are reported in Table 6. For all specifications, we find γ_{lGHG}^{UMC} to be negative and highly significant, consistent with our expectations. The one-factor model's coefficients imply that firms with a one standard deviation log-GHG emissions intensity above the cross-sectional mean have a negative exposure to unexpected changes in climate change concerns of about -0.024 (*i.e.*, the sum of the coefficients of *UMC* and $UMC \times lGHG$) in the non-asymmetric specification. We note that only the coefficient of the interaction between *UMC* and *lGHG* is significant across all three non-asymmetric specifications using different sets of controls. For the asymmetric specifications, firms with a one standard deviation log-GHG emissions intensity above the cross-sectional mean have an exposure of -0.045, and firms with a one standard deviation log-GHG emissions intensity below the cross-sectional mean have an exposure of 0.032. Results are similar for the other asymmetric specifications. We note that the common factor *UMC*, the interaction $UMC \times lGHG$ and the asymmetric term $UMC \times lGHG \times B$ are all significant across all of the three sets of factors. Overall, the results are in line with Pastor, Stambaugh, and Taylor (2020), Bolton and Kacperczyk (2020) and the results of our portfolio analysis.

[Insert Table 6 about here.]

4.3.2. Within-industry standardization of GHG

As noted by Ilhan, Sautner, and Vilkov (2020), most of the variation in GHG emissions intensity across firms can be attributed to industries. Bolton and Kacperczyk (2020) also find that institutional investors implement exclusionary screening based on direct emissions intensity in a few industries. We now test whether investors also consider the variation in GHG emissions intensity within industries when there are unexpected changes in climate change concerns. To do so, we re-estimate our panel regression models in (8) and (9), but now define $lGHG_{i,t}$ as the daily within-industry cross-sectionally standardized GHG emissions intensity of firm i at time t. The firms are grouped using the Fama-French 48 industry classification.[32]

Estimation results are reported in Table 7. As in our previous analyses, we find that the greener (browner) the firms are within an industry, the more positive (negative) their stock price's response is to unexpected changes in climate change concerns. However, we do not observe an asymmetry between firms that are above or below the within-industry average GHG emissions intensity. This result is expected, as Bolton and Kacperczyk (2020) suggest that institutional investors tend to screen firms on direct emissions intensity in a few salient industries. Thus, the asymmetry is only observed when comparing GHG emissions intensity across industries, not within industries. Moreover, the size and significance of coefficients is notably smaller than when we do not consider industry-specific greenness

[32]Industry data were retrieved from Kenneth French's website at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html`. We obtain similar results using a one- or two-digit SIC industry definition.

and standardize across all firms in our universe. While investors consider emissions intensity within industries, they put more emphasis on how firms compare to other firms generally, not necessarily to firms within their industry.

[Insert Table 7 about here.]

4.3.3. Firms that do not disclose GHG emissions

In our sample, we find that between 35.3% (in 2012) to 47.8% (in 2009) of firms do not disclose their GHG emissions; see Table 4. We can expect that these firms are also exposed to UMC. To test this, we use the industry average GHG as a proxy for the GHG emissions of the non-disclosing firms. For the panel of non-disclosing and disclosing firms, the generalized model becomes:

$$
\begin{aligned}
r_{i,t} = c &+ \left(\gamma^{lGHG} + \delta^{lGHG}_{A} A_{i,t} + \delta^{lGHG}_{UD} UD_{i,t} + \delta^{lGHG}_{A\text{-}UD} A_{i,t} UD_{i,t}\right) lGHG_{i,t} \\
&+ \Big(\gamma^{UMC} + \left(\gamma^{UMC}_{lGHG} + \delta^{UMC}_{lGHG\text{-}A} A_{i,t} + \delta^{UMC}_{lGHG\text{-}UD} UD_{i,t} + \delta^{UMC}_{lGHG\text{-}A\text{-}UD} A_{i,t} UD_{i,t}\right) lGHG_{i,t}\Big) UMC_t \\
&+ \boldsymbol{\beta}_i \mathbf{f}_t + \epsilon_{i,t}\,, \qquad (10)
\end{aligned}
$$

where $lGHG_{i,t}$ is defined as the industry average GHG emissions intensity for firms that do not disclose (and the reported emissions 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 coefficients of interest are $\delta^{UMC}_{lGHG\text{-}UD}$ and $\delta^{UMC}_{lGHG\text{-}A\text{-}UD}$, which measure the difference in exposure coefficients for the non-disclosing vs. disclosing firms.

Estimation results are reported in Table 8. We find that the difference of exposure coefficients are not significantly different from zero. This holds for all factor models considered. It follows that the returns of non-disclosing firms are exposed to the *UMC* factor in a similar way than returns of disclosing firms. This confirms that the prediction of Pastor, Stambaugh, and Taylor (2020) holds for all firms even if they do not disclose their GHG emissions.

[Insert Table 8 about here.]

5. Dimensions of climate change concerns

So far, we have established a relationship between unexpected changes in climate change concerns, proxied by the MCCC index and the *UMC* variable, and returns of green vs. brown firms. However, climate change is a broad subject with many facets, such as disasters, financial impacts, environmental impacts and regulatory impacts. Engle et al. (2020) suggest analyzing whether news about physical damages from climate change and news about regulatory risks have different impacts on stock returns. Moreover, the model of Pastor, Stambaugh, and Taylor (2020) implies that the effect of climate change concerns arises from two channels: (i) changes to expected cash flow and (ii) changes to investor tastes. Below, we build topical indices of Media Climate Change Concerns and analyze which dimensions drive the relationship between unexpected increases in climate change concerns and stock returns for green and brown firms. We then try to attribute these subjects to cash flow and/or taste channels.

To build the topical MCCC indices, we consider a topic-attribution weighted version of (3):

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

where $\theta_{k,n,t,s}$ is obtained from the estimated CTM (see Section 3.1 and Appendix A). We normalize and aggregate the scores for each index, following the steps of Section 2.3. This yields $K = 40$ topical MCCC indices.[33] We then estimate topical unexpected change in climate change concerns, which we denote by $UMC_{k,t}$, using the procedure outlined in Section 2.4 and Section 3.2. See Table 3 for the list of topics/themes.

To identify which themes and topics drive the relationship between climate change concerns and stock returns, we reconsider the approach of Section 4.2 with the following specification:

$$r_{p,t} = c_p + \beta_p^{UMC_k} UMC_{k,t} + \boldsymbol{\beta}_p \mathbf{f}_t + \epsilon_{p,t} \,. \tag{12}$$

[33] The MCCC index is obtained as a special case of (11) by setting $\theta_{k,n,t,s} = 1 \; \forall \; n, t, s$.

The quantity of interest is now $\beta_p^{UMC_k}$ rather than β_p^{UMC}.[34]

Regression results are reported in Table 9. We find that for all topics, the sign of the relationship is in line with our hypothesis: positive for the GMB and green portfolios, and negative for the brown portfolio. For 24 out of the 40 topics, the estimated coefficient is significant at the 5% confidence level.

Summarizing the results by theme, we find that "Financial and Regulation," "Agreement and Summit," "Societal Impact," "Research," and "Disaster" contain multiple topics with significant coefficients. Given the financial nature of the topics in "Financial and Regulation," we posit this theme primarily affects the cash flow channel. "Research", however, is likely to only affect the tastes channel, as research results hardly impact firms' cash flows, at least in the short-term. The "Agreement and Summit" theme is likely to affect both channels. On the one hand, regulations can have direct impacts on firms' future cash flows. On the other hand, the discussions taking place at these conferences often underline the future disastrous consequences of climate change, which can affect investors' tastes. We posit that the "Societal Impact" theme is likely to affect the tastes channel. However, because of Topic 30, which discusses funding for green programs, the societal impact theme could also affect the cash flow channel through subsidies to green firms and green projects. Finally, we believe "Disaster" primarily affects the taste channel by emphasizing the direct impacts of climate change if strong actions are not taken.

[Insert Table 9 about here.]

6. Conclusion

Our paper empirically verifies the prediction of Pastor, Stambaugh, and Taylor (2020) 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

[34] The analysis based on the panel specification presented in Section 4.3 yields the same conclusion as (12).

change from major U.S. newspapers 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) from the prediction error of a first-order autoregressive model.

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. We also document that the size of the effect is larger in absolute value for brown stocks than for green stocks. This result is consistent with institutional investors using screening methods to disinvest in browns stock and reinvest in the rest of the market (*i.e.*, not only in green stocks).

Finally, we construct topical MCCC indices to determine whether the stock price reaction to UMC arises from expectations about a firm's cash flow or changes to investors' preferences, as suggested by Pastor, Stambaugh, and Taylor (2020). We estimate a correlated topic model on our news corpus. Our analysis identifies eight themes (*i.e.*, clusters of topics), of which five are significantly related to the stock returns of green vs. brown firms. Our results suggest that the effect arises from changes to both cash flow expectations and investor tastes.

A key message for business leaders is that climate change concerns also matter for their firms' equity value 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 (see 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

Amihud, Y. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5:31–56.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang. 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61:259–99.

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

Bali, T. G., S. J. Brown, and Y. Tang. 2017. Is economic uncertainty priced in the cross-section of stock returns? *Journal of Financial Economics* 126:471–89.

Bali, T. G., N. Cakici, and R. F. Whitelaw. 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics* 99:427–46.

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

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.

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. 2020. Do investors care about carbon risk? Working Paper, NBER.

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.

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

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.

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.

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

———. 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.

Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81:607–36.

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.

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.

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

GSIA. 2018. Global Sustainable Investment Review 2018. Working Paper.

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.

Harvey, C. R., Y. Liu, and H. Zhu. 2016. ... and the cross-section of expected returns. *Review of Financial Studies* 29:5–68.

Harvey, C. R., and A. Siddique. 2000. Conditional skewness in asset pricing tests. *Journal of Finance* 55:1263–95.

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

Hou, K., C. Xue, and L. Zhang. 2015. Digesting anomalies: An investment approach. *Review of Financial Studies* 28:650–705.

Ilhan, E., Z. Sautner, and G. Vilkov. 2020. Carbon tail risk. *Review of Financial Studies* Forthcoming.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48:65–91.

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

Larsen, V. 2017. Components of uncertainty. Working Paper.

Larsen, V., and L. A. Thorsrud. 2017. Asset returns, news topics, and media effects. Working Paper.

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., and M. Lee. 2014. Low-dimensional embeddings for interpretable anchor-based topic inference. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1319–28.

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.

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

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

Ramelli, S., A. F. Wagner, R. J. Zeckhauser, and A. Ziegler. 2018. Investor rewards to climate responsibility: Evidence from the 2016 climate policy shock. Working Paper.

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.

Schwarz, G., et al. 1978. Estimating the dimension of a model. *Annals of Statistics* 6:461–4.

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: Sources of climate change news

This table reports, for each source, the number of articles discussing climate change, the total number of articles published, and the percentage of articles that address climate change from January 2003 to June 2018.

Source	Climate	Total	%
Wall Street Journal	3,776	1,673,007	0.23
New York Times	3,711	1,477,936	0.25
Washington Post	2,323	1,029,917	0.23
Los Angeles Times	1,594	747,557	0.21
Chicago Tribune	509	1,058,643	0.05
USA Today	249	149,450	0.17
New York Daily News	129	220,002	0.06
New York Post	109	190,880	0.06

Table 2: Most concerning climate change articles
This table reports the most concerning news articles according to our article-level concerns score. We report the publication date, the concerns score, the negativity level, the level of risk, the first 50 characters of the article's headline (in the original format), and the source. The weight is defined as the level of risk multiplied by the level of negativity. The level of risk and negativity are computed using the LIWC2015 sentiment and risk lexicons.

Rank	Date	Concerns	Risk	Negativity	Title	Source
1	2014-06-25	4.80	5.48	0.88	Warning on biz 'climate'	New York Daily News
2	2018-01-05	3.44	3.98	0.87	Storms, Fires and Floods Lead to Record Payouts	New York Times
3	2018-01-04	3.44	3.98	0.86	2017 Set a Record for Losses From Natural Disaster	New York Times
4	2010-03-12	3.04	3.04	1.00	The Climate Rules Weren't From Us	Wall Street Journal
5	2018-02-05	3.02	3.92	0.77	An urgent national security issue	Washington Post
6	2011-12-08	2.89	3.55	0.82	2011 saw record number of high-cost weather disast	Los Angeles Times
7	2017-02-22	2.79	4.19	0.67	Questions for: 'Mexico City, Parched and Sinking,	New York Times
8	2011-12-09	2.77	3.69	0.75	A dozen billion-dollar weather disasters in 2011	Chicago Tribune
9	2018-06-23	2.76	3.45	0.80	Picturing climate change	Washington Post
10	2008-10-07	2.73	2.88	0.95	One in 4 Mammals Threatened With Extinction, Group	New York Times
11	2015-05-21	2.72	5.07	0.54	Global warming is risk to security, Obama says	Chicago Tribune
12	2005-03-12	2.66	3.30	0.81	'I Have a Nightmare'	New York Times
13	2014-10-22	2.63	3.47	0.76	A Retreat From Weather Disasters	New York Times
14	2008-01-03	2.60	3.85	0.67	Disasters the 'new normal'	USA Today
15	2016-06-19	2.59	3.46	0.75	EMP: More Urgent Than Any Climate Threat; It is in	Wall Street Journal
16	2018-01-26	2.59	3.15	0.82	Doomsday Clock moves 30 seconds ahead, landing at	Washington Post
17	2014-04-06	2.57	3.80	0.68	The perils of climate change	Washington Post
18	2017-08-09	2.55	3.94	0.65	Don't Ignore the Dire Threat of Climate Change	New York Times
19	2018-03-27	2.51	2.61	0.96	Degraded land drives migration, report says	Washington Post
20	2015-05-21	2.49	4.50	0.55	Obama warns of climate threat to U.S. security	Los Angeles Times

Table 3: Highest probability words for each topic

This table reports, for each topic k, the words with the ten highest probability $\omega_{v,k}$. The topics are estimated using the Correlated Topic Model. We estimate $K = 40$ topics. For ease of interpretation, we split the topics into eight themes.

Theme	Topic	1	2	3	4	5	6	7	8	9	10
Financial and Regulation	Topic 40	project	technology	plant	cost	coal	carbon_dioxide	power_plant	facility	scale	carbon
	Topic 32	car	vehicle	standard	truck	automaker	diesel	emission	auto	engine	fuel
	Topic 31	oil	tax	fuel	price	carbon_tax	production	taxis	cost	ethanol	revenue
	Topic 25	home	business	product	consumer	building	panel	energy_efficiency	customer	bulb	light
	Topic 21	market	industry	emission	permit	credit	system	allowance	cap	price	cost
	Topic 17	investor	investment	business	executive	risk	firm	fund	bank	shareholder	asset
	Topic 16	bill	legislation	vote	measure	lawmaker	senator	governor	proposal	sen	gov
	Topic 15	power	electricity	coal	plant	wind	utility	capacity	power_plant	reactor	renewable
	Topic 13	gas	methane	chemical	leak	waste	ozone	production	industry	carbon_dioxide	atmosphere
	Topic 7	airline	flight	air	aviation	airport	pollution	plane	aircraft	travel	emission
	Topic 6	rule	administration	agency	regulation	law	court	decision	authority	administrator	action
Agreement and Summit	Topic 37	leader	article	pope	trade	security	official	meeting	summit	trump	visit
	Topic 35	obama	campaign	trump	election	candidate	voter	party	policy	position	job
	Topic 19	pipeline	mr_obama	mr_trump	coal	job	project	decision	land	oil	mine
	Topic 18	agreement	deal	talk	meeting	commitment	conference	target	accord	treaty	official
	Topic 14	email	science	document	headline	information	investigation	research	letter	statement	committee
Societal Impact	Topic 38	policy	action	cost	solution	economy	planet	reason	future	growth	politician
	Topic 34	poll	survey	majority	public	pew	penguin	concern	opinion	result	support
	Topic 30	money	program	budget	development	fund	funding	effort	initiative	aid	poverty
	Topic 11	child	school	student	family	woman	life	street	art	event	police
	Topic 9	health	death	disease	security	population	threat	child	life	war	risk
	Topic 8	science	book	story	truth	film	news	movie	medium	reader	life
Research	Topic 22	earth	atmosphere	planet	space	cloud	science	satellite	system	research	sun
	Topic 5	study	researcher	temperature	research	paper	finding	effect	change	author	activity
	Topic 3	datum	temperature	model	trend	record	estimate	period	figure	increase	rate
Disaster	Topic 36	weather	storm	hurricane	record	temperature	event	heat	heat_wave	drought	wind
	Topic 33	fire	wildfire	insurance	risk	home	property	disaster	loss	flood	zone
	Topic 24	city	mayor	building	resident	community	plan	county	housing	neighborhood	official
	Topic 12	island	sea	sea_level	storm	flood	flooding	land	beach	home	village
Environmental Impact	Topic 39	ice	glacier	snow	ice_sheet	ocean	temperature	sea_level	satellite	mountain	researcher
	Topic 29	team	researcher	lake	rock	period	study	layer	sample	evidence	eruption
	Topic 28	forest	tree	land	deforestation	carbon	plant	wood	soil	rain_forest	pine
	Topic 10	specie	bird	park	habitat	animal	wildlife	conservation	population	extinction	plant
	Topic 1	ship	drilling	oil	sea	fishing	shipping	coast	boat	shell	exploration
Agricultural Impact	Topic 23	farmer	crop	farm	agriculture	plant	soil	corn	rice	land	wheat
	Topic 20	food	animal	meat	cow	cattle	farm	ski	resort	beef	diet
	Topic 4	drought	region	river	rain	desert	lake	dam	rainfall	water_supply	mountain
	Topic 2	wine	grape	coffee	region	fruit	vineyard	temperature	sugar	bottle	harvest
Other	Topic 27	bear	permafrost	ice	sea_ice	population	seal	species	hunting	animal	wildlife
	Topic 26	reef	ocean	coral	algae	fish	barrier_reef	sea	ecosystem	acidification	tourism

Rank

Probability: 0.05, 0.10

Table 4: Summary statistics of the revenue-scaled greenhouse gas emissions variable

This table reports summary statistics of the revenue-scaled greenhouse gas emissions level used to establish firms' greenness and brownness. Panel A reports the percentage of firms n the S&P 500 universe with available greenhouse gas emissions data for each year. Panel B reports summary statistics.

Panel A: Percentage of firms with emissions data	
Year	*GHG*
2009	52.27
2010	57.44
2011	62.92
2012	64.73
2013	58.61
2014	57.30
2015	56.50
2016	60.27
2017	60.14
Panel B: Summary statistics	
Statistics	*GHG*
Observations	2,429
Average	680.57
Standard deviation	1,584.01
Skewness	3.22
Kurtosis	10.74
Minimum	1.10
25th percentile	21.41
Median	59.52
75th percentile	368.41
Maximum	9,445.71

Table 5: Regression results of portfolios' returns

This table reports regression results of the daily unexpected changes in climate change concerns, *UMC*, and control factors on the returns of GMG, green, brown and neutral portfolios; see Eq. (7). The composition of the four portfolios is based on firms' annual revenue-scaled GHG emissions. The model is estimated with data from January 2010 to June 2018. 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% level, respectively.

	GMB	Green	Brown	Neutral
Intercept	0.000	0.001	0.001	0.005
	(0.011)	(0.005)	(0.008)	(0.003)
UMC	0.090^{***}	0.037^{***}	-0.054^{***}	0.022^{***}
	(0.027)	(0.012)	(0.019)	(0.008)
MKT	0.161^{***}	1.105^{***}	0.944^{***}	1.033^{***}
	(0.017)	(0.007)	(0.014)	(0.005)
HML	0.163^{***}	0.184^{***}	0.021	-0.090^{***}
	(0.046)	(0.022)	(0.03)	(0.013)
SMB	0.121^{***}	0.025^{**}	-0.097^{***}	0.020^{**}
	(0.034)	(0.012)	(0.027)	(0.009)
CMA	-0.559^{***}	-0.103^{***}	0.456^{***}	0.235^{***}
	(0.063)	(0.031)	(0.046)	(0.017)
RMW	-0.209^{***}	-0.109^{***}	0.100^{**}	0.150^{***}
	(0.048)	(0.019)	(0.041)	(0.015)
MOM	0.133^{***}	-0.061^{***}	-0.194^{***}	-0.075^{***}
	(0.027)	(0.012)	(0.022)	(0.007)

Table 6: Panel regression results of individual firms' returns

This table reports panel regression results about the effect of the daily standardized logarithmic revenue-scaled GHG emissions (intensity) on stock-level exposure to unexpected changes in climate change concerns, *UMC*. The regression is estimated using data on S&P 500 firms from January 2010 to June 2018. We report the intercept and the exposure to *lGHG* (*i.e.*, γ^{lGHG}), the exposure coefficients to the unexpected changes in climate change (*i.e.*, γ^{UMC}) and *lGHG*×*UMC* (*i.e.*, γ^{UMC}_{lGHG}); see Eq. (8). We also report the coefficients for the asymmetric effect *lGHG*×*A* (*i.e.*, δ^{lGHG}_{A}) and *lGHG*×*A*×*UMC* (*i.e.*, $\delta^{UMC}_{lGHG\text{-}A}$); see Eq. (9). For the controls, we use one-factor (*i.e.*, *MKT*), three-factor (*i.e.*, *MKT*, *HML*, *SMB*) and six-factor (*i.e.*, *MKT*, *HML*, *SMB*, *RMW*, *CMA*, *MOM*) models. Standard errors of the estimators are reported in parentheses. The signs *, **, and *** indicate significant coefficients at the 10%, 5%, and 1% level, respectively.

	One-factor		Three-factor		Six-factor	
	Eq. (8)	*Eq.* (9)	*Eq.* (8)	*Eq.* (9)	*Eq.* (8)	*Eq.* (9)
Intercept	0.001	0.002	0.003^{**}	0.004^{**}	0.002	0.003^{*}
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
lGHG	0.001	0.002	0.001	0.002	−0.001	0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
lGHG×*A*		−0.004		−0.003		−0.003
		(0.004)		(0.004)		(0.004)
UMC	0.007^{**}	0.018^{***}	0.004	0.014^{***}	0.006^{**}	0.015^{***}
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
lGHG×*UMC*	-0.031^{***}	-0.014^{**}	-0.029^{***}	-0.013^{**}	-0.027^{***}	-0.013^{**}
	(0.004)	(0.006)	(0.004)	(0.006)	(0.004)	(0.006)
lGHG×*A*×*UMC*		-0.049^{***}		-0.045^{***}		-0.041^{***}
		(0.011)		(0.011)		(0.011)

Table 7: Panel regression results of individual firms' returns: Within-industry effect
This table reports panel regression results for the effect of the daily standardized logarithmic revenue-scaled GHG emissions (intensity) on stock-level exposure to unexpected changes in climate change concerns, *UMC*. See Table 6 for more details.

	One-factor		Three-factor		Six-factor	
	Eq. (8)	*Eq.* (9)	*Eq.* (8)	*Eq.* (9)	*Eq.* (8)	*Eq.* (9)
Intercept	0.001	0.002	0.003^{**}	0.003^{*}	0.002	0.002
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
$lGHG$	0.001	0.001	0.001	0.001	0.000	0.000
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
$lGHG\times A$		0.000		0.000		0.000
		(0.002)		(0.002)		(0.002)
UMC	0.006^{*}	0.012^{***}	0.003	0.008^{*}	0.005^{*}	0.010^{**}
	(0.003)	(0.004)	(0.003)	(0.004)	(0.003)	(0.004)
$lGHG\times UMC$	-0.017^{***}	-0.011^{***}	-0.017^{***}	-0.011^{***}	-0.016^{***}	-0.010^{***}
	(0.002)	(0.004)	(0.002)	(0.004)	(0.002)	(0.004)
$lGHG\times A\times UMC$		-0.011^{**}		-0.009^{*}		-0.009^{*}
		(0.006)		(0.005)		(0.005)

Table 8: Panel regression results: Non-disclosure effect

This table reports panel regression results about the effect of daily and standardized logarithm revenue-scaled GHG emissions (intensity) on stock-level exposure to unexpected changes in climate change concerns, *UMC*. The standardized GHG emissions intensity of firms that do not disclose their GHG emissions level is set at the average for the firm's industry. *UD* takes a value of one when emissions data is not disclosed; see Eq. (10). See Table 6 for more details.

	One-factor	Three-factor	Six-factor
Intercept	0.002	0.003^{**}	0.003^{*}
	(0.002)	(0.002)	(0.002)
$lGHG$	0.003	0.002	0.000
	(0.002)	(0.002)	(0.002)
$lGHG\times A$	-0.003	-0.003	-0.003
	(0.005)	(0.004)	(0.004)
$lGHG\times UD$	-0.013^{*}	-0.012	-0.010
	(0.008)	(0.008)	(0.008)
$lGHG\times A\times UD$	0.007	0.005	0.003
	(0.012)	(0.012)	(0.012)
UMC	0.018^{***}	0.014^{***}	0.015^{***}
	(0.004)	(0.004)	(0.004)
$lGHG\times UMC$	-0.014^{**}	-0.013^{**}	-0.012^{**}
	(0.006)	(0.006)	(0.006)
$lGHG\times A\times UMC$	-0.046^{***}	-0.043^{***}	-0.040^{***}
	(0.011)	(0.011)	(0.011)
$lGHG\times UD\times UMC$	0.010	0.011	0.009
	(0.019)	(0.019)	(0.018)
$lGHG\times A\times UD\times UMC$	-0.039	-0.036	-0.025
	(0.030)	(0.029)	(0.029)

Table 9: Regression results of portfolios' returns and the topical MCCC indices

This table reports the estimates of $\beta_{k,p}^{UMC}$ in the regression of topical daily unexpected changes in climate change concerns on GMB, green, brown and neutral portfolios; see Eq. (12). Rows indicate which topic the MCCC index is based on, while columns indicate which portfolio's returns are being analyzed. The signs *, **, and *** indicate significant coefficients at the 10%, 5%, and 1% level, respectively. They are obtained via t-stats for which the standard error of the estimator is estimated using Newey and West (1987, 1994). The regressions are estimated with data from January 2010 to June 2018.

	GMB	Green	Brown	Neutral
Theme "Financial and Regulation"				
Topic 40	0.089**	0.043***	−0.046*	0.008
Topic 32	0.054*	0.035**	−0.019	0.017
Topic 31	0.080**	0.035**	−0.045*	0.012
Topic 25	0.080**	0.032*	−0.047*	0.019*
Topic 21	0.132***	0.056***	−0.076**	0.014
Topic 17	0.062**	0.023*	−0.038*	0.003
Topic 16	0.093***	0.032**	−0.061**	0.019**
Topic 15	0.066**	0.028**	−0.038**	0.005
Topic 13	0.024	0.028**	0.004	0.021***
Topic 7	0.092***	0.041***	−0.051***	0.026***
Topic 6	0.045	0.016	−0.030	0.020**
Theme "Agreement and Summit"				
Topic 37	0.067***	0.015	−0.052***	0.005
Topic 35	0.063***	0.019**	−0.043***	0.014**
Topic 19	0.036**	0.009	−0.027**	0.005
Topic 18	0.119***	0.036**	−0.082***	0.016
Topic 14	0.051*	0.016	−0.035*	0.015*
Theme "Societal Impact"				
Topic 38	0.068**	0.029**	−0.039*	0.025***
Topic 34	0.056**	0.023*	−0.032*	0.022***
Topic 30	0.106***	0.041***	−0.066***	0.023***
Topic 11	0.055**	0.029**	−0.026	0.016*
Topic 9	0.041	0.024*	−0.017	0.009
Topic 8	0.039	0.011	−0.028	0.023**
Theme "Research"				
Topic 22	0.107***	0.051***	−0.057**	0.035***
Topic 5	0.045	0.027**	−0.017	0.018**
Topic 3	0.078**	0.038***	−0.040*	0.018**
Theme "Disaster"				
Topic 36	0.027	0.014	−0.013	0.016**
Topic 33	0.057**	0.019	−0.037**	0.003
Topic 24	0.044*	0.021**	−0.023	0.017**
Topic 12	0.053*	0.022*	−0.030	0.016**
Theme "Environmental Impact"				
Topic 39	0.010	0.012	0.001	0.011
Topic 29	0.047	0.023*	−0.024	0.011
Topic 28	0.076*	0.045**	−0.032	0.012
Topic 10	0.048	0.020	−0.028	0.009
Topic 1	0.037	0.023	−0.013	0.020**
Theme "Agricultural Impact"				
Topic 23	0.008	0.019	0.011	0.008
Topic 20	0.007	0.010	0.003	0.006
Topic 4	0.034	0.025**	−0.009	0.010
Topic 2	0.013	0.017	0.004	0.023**
Theme "Other "				
Topic 27	0.065	0.016	−0.049	0.009
Topic 26	0.028	0.024	−0.004	0.019**

Figure 1: Number of article equivalents by theme
This figure displays the monthly number of article-equivalent publications for each theme from January 2010 to June 2018.

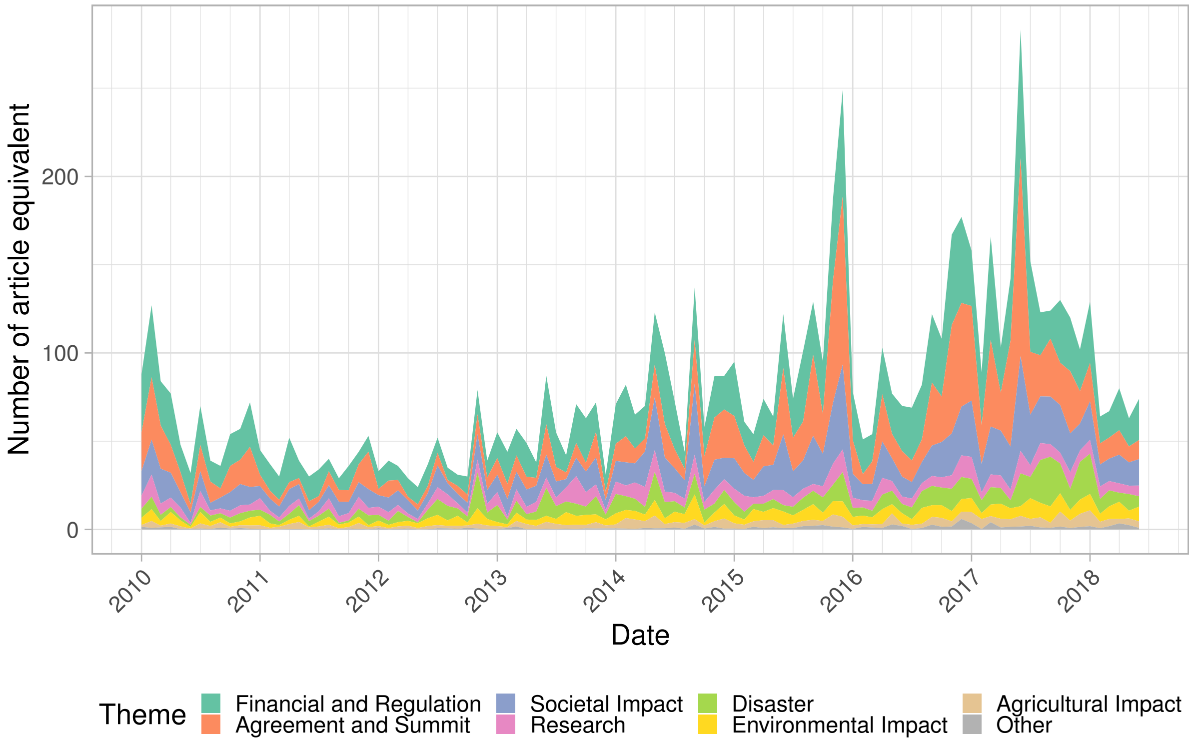


Figure 2: Media Climate Change Concerns index

This figure displays the daily MCCC index (gray points) together with its 30-day moving average (bold line) for 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 to be forward-looking, since the data from that period is used to compute the source-specific standard deviation estimate necessary 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.

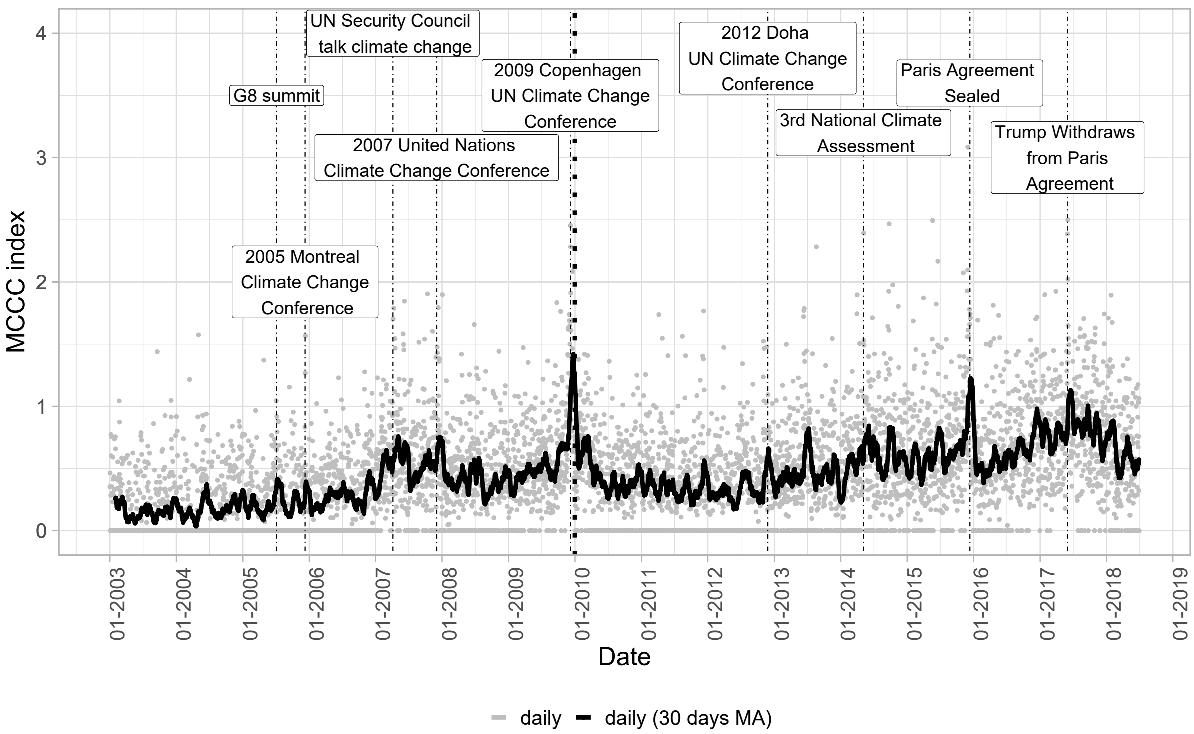
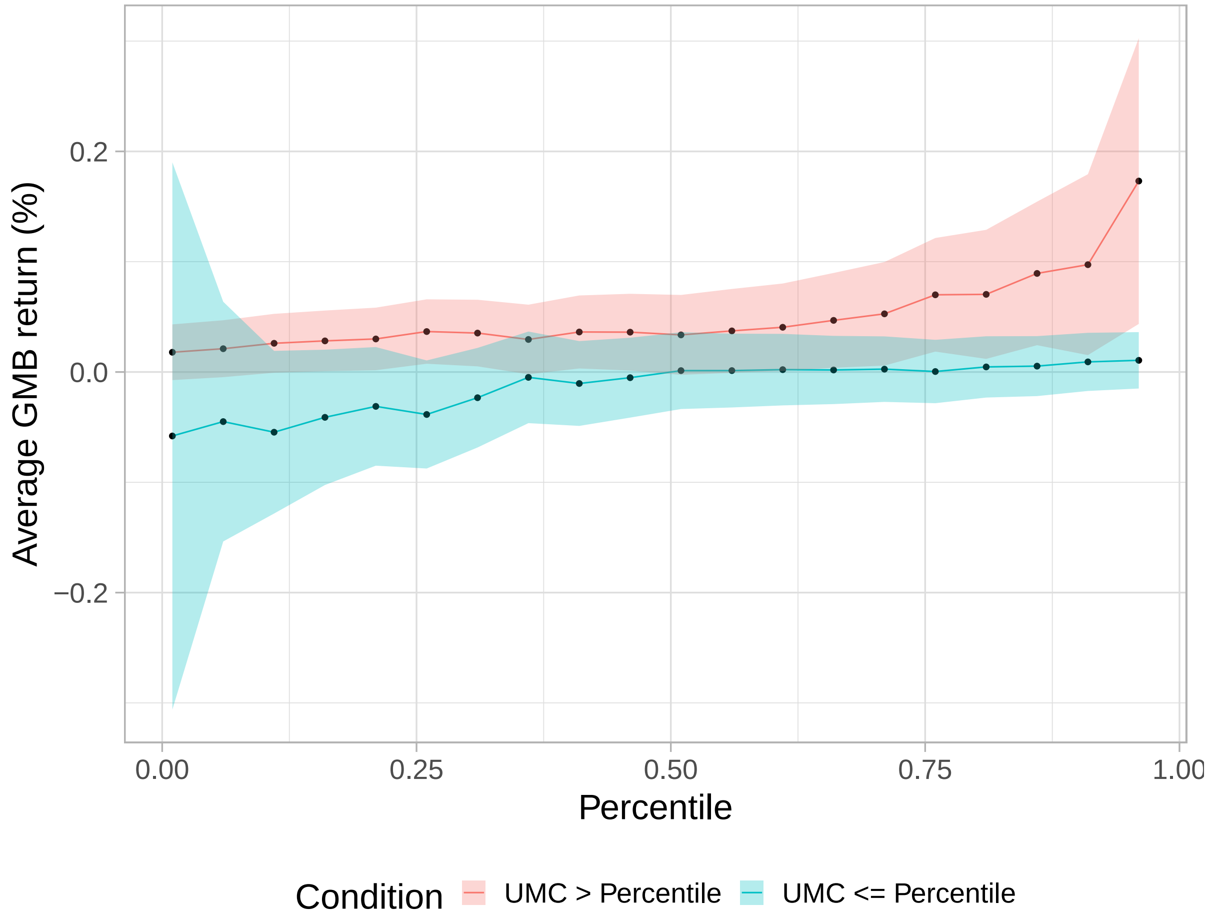


Figure 3: Green minus brown portfolio average return
This figure displays the average return of the GMB portfolio (vertical axis) conditional on *UMC* being above or below a specific threshold (horizontal axis). Thresholds are set as percentiles of *UMC*. The colored bands report the 95% confidence interval.



Appendix A. Topic modeling and theme construction

To improve the estimation of the topic model, 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 the ones that appear more than 100 times in the climate change corpus. An example of such a collocation is "climate change."

Next, we lemmatize every standalone word (*i.e.*, excluding collocations). That is, we use vocabulary and morphological analysis of words to remove inflectional endings and transform words 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).

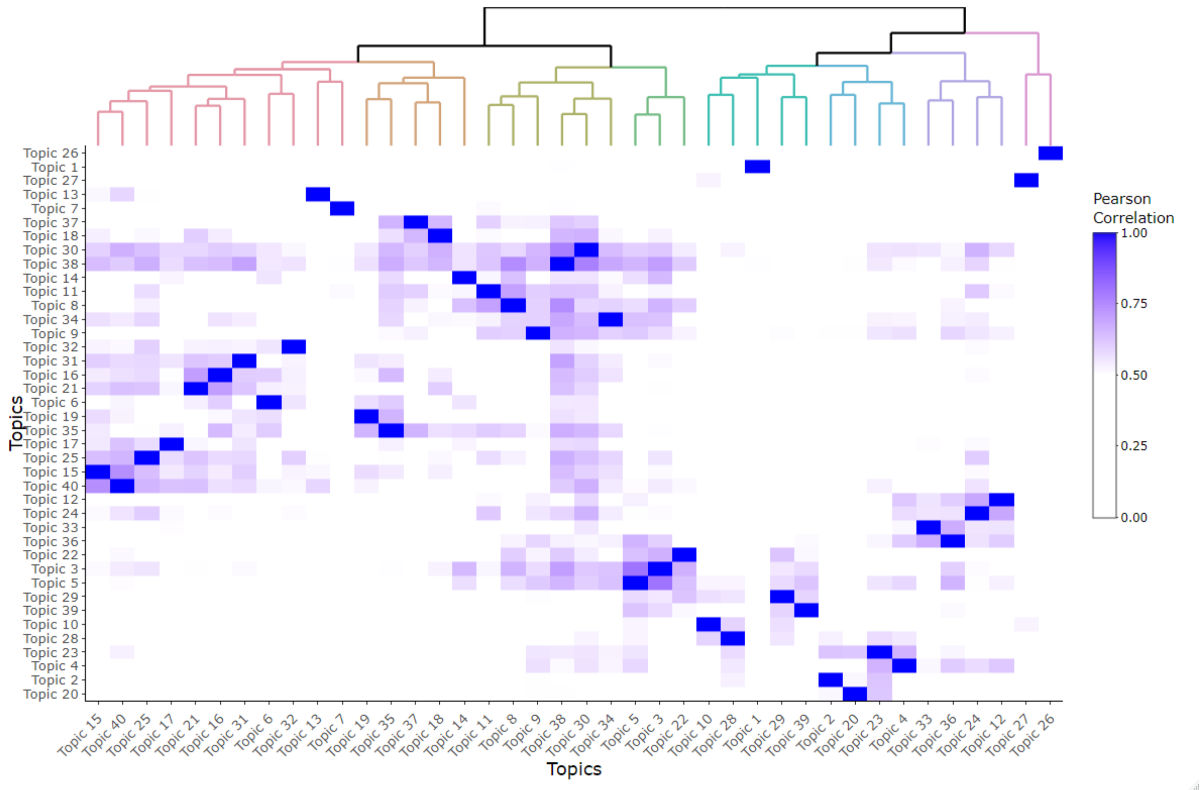
Following Hansen, McMahon, and Prat (2018), we estimate $K = 40$ topics.[35] This is a good balance between too few topics, which tend to be overly general, and too many topics, which can be too specific. We use the R package **STM** from Roberts, Stewart, and Tingley (2014) to estimate the correlated topic model.

To construct themes, we begin by computing the correlation matrix of the topical MCCC indices. Then, we perform hierarchical clustering on this correlation matrix, where we settle on eight clusters for interpretability purposes. In Figure A.1, we display the correlation matrix as well as the dendogram generated from the hierarchical clustering algorithm, where correlations below 0.5 are kept blank for better visualization of the clusters. The correlation matrix is also reordered to put clusters side-to-side.

[35] Using the approach of Mimno and Lee (2014), we find that the optimal number of topics is 75. Hansen, McMahon, and Prat (2018) estimate the optimal number of topics at 70, but reduce this to 40 for ease of interpretation. We follow their reasoning.

Figure A.1: Correlation matrix of the topical MCCC indices

This figure displays the correlation of the topical MCCC indices. The correlation matrix is rearranged according to a hierarchical clustering algorithms to highlight clusters. Correlations below 0.5 are kept blank. The colors in the dendrogram tree highlight the set of 8 clusters obtained with the clustering algorithm.



Appendix B. Fama-MacBeth regression test

As an additional test, we consider a multivariate Fama-MacBeth cross-sectional regression framework (Fama and MacBeth 1973):

$$r_{i,t} = c_t + \lambda_t^{lGHG} lGHG_{i,t} + \boldsymbol{\lambda}_t \mathbf{CTRL}_{i,t} + \epsilon_{i,t}\,, \tag{B.1}$$

where our focus is now on λ_t^{lGHG}, which represents the effect of log-GHG emissions intensity on stock returns at each point in time. Thus, if λ_t^{lGHG} is positive (negative), brown firms have a higher (lower) return than green firms on day t, controlling for other firms' characteristics considered in $\mathbf{CTRL}_{i,t}$. We can capture the contemporaneous relationship between that GHG emissions intensity effect and UMC as follows:

$$\lambda_t^{lGHG} = c + \beta\, UMC_t + \eta_t\,. \tag{B.2}$$

Under the model of Pastor, Stambaugh, and Taylor (2020), returns of high (low) GHG emissions intensity firms are lower (higher) when there is an unexpected increase in climate change concerns. Thus, we expect β to be negative.

We consider several firms' characteristics in the controls $\mathbf{CTRL}_{i,t}$: the stock's market beta, size, and book-to-market ratio (Fama and French, 1992); the momentum and reversal (Jegadeesh and Titman, 1993); the stock's monthly co-skewness (Harvey and Siddique, 2000); the stock's illiquidity measure (Amihud, 2002); the idiosyncratic volatility and stock exposure to aggregate stock market volatility (Ang et al., 2006), the annual growth rate of total assets and quarterly returns on equity (Hou, Xue, and Zhang, 2015); and the lottery-like stock characteristic (Bali, Cakici, and Whitelaw, 2011). We refer the reader to (Bali, Brown, and Tang, 2017) for details on the computation of these variables. We first proceed by estimating λ_t^{lGHG} in (B.1) cross-sectionally for each day using various sets of control characteristics.

In Table B.1, we report the estimations of the β in (B.2) for the various specifications of controls. For all cases, we find a negative and significant coefficient β, consistent with the hypothesis that the stock returns of brown (green) firms are reduced (increased) when there are unexpected increases in climate change concerns.

Table B.1: Fama-MacBeth regression results

This table reports the estimated intercept and exposure of the logarithm of the GHG emissions intensity coefficient to the *UMC*; see Eq. (B.2). The daily GHG emissions intensity coefficient is first estimated by running a cross-sectional regression for each day, controlling for firm-level characteristics; see Eq. (B.1). We consider six specifications, each consisting of various sets of firm-characteristic variables for the controls: Specification (1) does not include any controls; (2) includes the market beta characteristic; (3) extends the set in (2) to control for exposure to aggregate volatility; (4) extends the set in (3) to control for the size, book-to-market ratio and momentum characteristics; (5) extends (4) to control for reversal, illiquidity, coskewness, idiosyncratic volatility, annual growth on assets and return on equity characteristics. Finally, (6) extends (5) with the lottery-like characteristic. We refer to Bali, Brown, and Tang (2017) for details on how to construct the characteristics. 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% level, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-0.002	-0.002	-0.002	-0.002	-0.002	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
UMC	-0.018^{***}	-0.016^{***}	-0.015^{***}	-0.012^{***}	-0.010^{**}	-0.009^{**}
	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)