Alternative Thinking | Q4 2021

Can Machine Learning Help Manage Climate Risks?

Executive Summary

As climate change is becoming an increasingly observed phenomenon and understood to be caused by carbon and carbon-equivalent emissions, investors are adjusting their portfolios to prepare for a future regime shift to a lower-carbon economy. The primary approach taken to prepare for this change is to incorporate carbon emissions in investment selection. However, albeit a practical approach, carbon emissions can be a narrow measure of overall climate risk.

In this article we explore the insights of Engle, Giglio, Kelly, Lee, and Stroebel in "Hedging Climate Change News", 2020, where they use textual analysis and machine learning techniques to create a broad climate hedging portfolio based on stocks' sensitivity to climate news. We find that, subject to further research, these insights could be used as a complement to carbon-aware investing in defending against climate change.

PSG
Portfolio Solutions Group

Contents

About the Portfolio Solutions Group

The Portfolio Solutions Group (PSG) provides thought leadership to the broader investment community and custom analyses to help AQR clients achieve better portfolio outcomes.

We thank Bryan Kelly, Pete Hecht and Alfie Brixton for their work on this paper. We also thank Lukasz Pomorski for his helpful comments.

Introduction: The Benefits, and Limits, of Carbon Awareness

The topic of climate-aware investing (and ESG investing more broadly) is now widespread in the asset management industry. The primary practical action investors take to address climate risk in their portfolio is tilting holdings away from high carbon-emitting firms towards those with lower emissions, as discussed by our colleagues Palazzolo, Pomorski, and Zhao in "(Car)Bon Voyage: The Road to Low Carbon Investment Portfolios" (2021). Our colleagues highlight risk management as one motivation for these carbon emissions-based tilts. For example, as public and private institutions take steps to stave off the adverse effects of climate change on our collective well-being, there is increasing probability that the economy will transition to less carbon-intensive means of production. This is a risk to incumbent firms that are especially carbon dependent (barring significant change in their operations), and an opportunity for existing firms and new entrants with a smaller carbon footprint that might be more emission-productive. As the world's awareness of climate risk grows, and institutions mobilize to counteract this risk, the prudent investor may decide to hedge it by overweighting assets that are likely to thrive in a new carbon-light economy.

But why is the investment community focused on the idea of implementing climate-related portfolio objectives based on carbon emissions? Climate change and carbon emissions are linked, but not synonymous. Part of this emphasis is rooted in climate science and research links between atmospheric carbon, environmental change, and economic outcomes (e.g., Nordhaus, 2014). But focusing on carbon emissions is a practical choice as well; there are few observable indicators beyond carbon emissions that immediately link firms' behavior to climate change. However, emissions data are noisy reflections of the true climate exposure of a firm. This is both because some emissions (e.g., scope 3) may be difficult to measure, and important drivers of climate exposure may not be captured by emissions (e.g., fossil fuel reserves still in the ground, physical climate change exposure, etc.).

In this article, we revisit the work of Engle, Giglio, Kelly, Lee, and Stroebel in "Hedging Climate Change News" (2020) and lay out a framework for using machine learning tools to help manage climate risk. We focus on climate-aware investing from a risk management perspective and take a novel angle on measurement of the climate-risk sensitivity of stocks using textual analysis and machine learning. The ultimate result is a distinctive climate hedge portfolio that can complement the reduced-carbon portfolio approach of Palazzolo, Pomorski, and Zhao, 2021, as part of an investor's broad climate risk reduction strategy. We emphasize that we do not present the only, or necessarily optimal approach to this problem, but rather a toolkit which investors can use as a base or guide to tackling this kind of problem. This is an area of active research for us where we plan to incrementally improve on the methodology proposed herein.

Challenges of Climate Hedging

The risk of climate change and the increasing attention it receives among financial market participants prompts two critical asset management questions. First, can climate risk have a material impact on the value of a portfolio? Second, how can investors insure their savings against the adverse effects of climate change?

Answering these questions requires overcoming difficult **measurement** challenges that stem from at least three factors:

- First, the rise in investor attention on climate change is a fairly recent phenomenon. The measurable impact of climate change on asset markets is therefore expected to manifest only in recent data. This means we may only have short samples to detect any effects.
- Second, while the adverse effects of climate change may have started to materialize in the form of extreme weather and other natural disasters, the most severe consequences are expected to occur in the distant future. The net present value of expected climate consequences may be surprisingly small today due to very long-dated discounting – despite those consequences being potentially catastrophic in future-value terms. As a result, it may be difficult to distinguish climate effects in current asset prices whose valuations are dominated by nearer-term cash flow considerations such as productivity growth, political and regulatory uncertainty, or the already catastrophic COVID-19 pandemic.
- Third, but not least, the effects of climate change are shrouded by an extraordinary degree of uncertainty. This comes not only in the form of "known" risks (e.g., a well-defined standard deviation in long-run temperatures in a given climate model), but also significant ambiguity about climate change (e.g., uncertainty about which is the correct model of long-run temperatures, or uncertainty over policy actions which could affect asset valuations).

In addition to the measurement challenges, there are contracting or **financial engineering** challenges. Naturally, an ideal climate-hedge solution would be a market where investors could trade claims explicitly linked to long-dated climate outcomes. At least two factors cast doubt on the feasibility of such a solution:

- The first is the question of "what to hedge," i.e., what is the outcome insurance contracts should protect against? It is difficult to predict what the effects of climate change will be, and even more so the expected scale of damages.
- Second, even if market participants could agree on some sort of climate index to link with derivatives, we are talking about a long-dated, aggregate, non-diversifiable risk. History tells us that such insurance contracts are extremely difficult to implement. For example, Nobel laureate Robert Shiller's idea of using derivatives on house price indices to hedge home equity risk has proven difficult in practice (Sommervoll and Swidler, 2021). As put by the Governor of the Bank of England, Andrew Bailey, "we cannot diversify away from our exposure to the planet."

Insights From Option-Pricing Theory

We find guidance for solving the difficult problem of climate risk hedging in the seminal theory of option prices. Recall that a put option is simply an insurance contract—it pays off a dollar for every dollar that the reference asset drops below some "deductible" value. The revolutionary (and Nobel-prize-garnering) insight of Black-Scholes (1973) and Merton (1973) is that the fair value of the insurance contract can be inferred by a *replication* argument. Without delving into the technical details, the intuition of the replication argument is that an investor can replicate the insurance payout simply by holding the reference asset in a clever and dynamic way. Naturally, the value of the insurance contract and the reference asset are inextricably linked.

For example, suppose stock "A" is trading today at $100. In the options market, there is an insurance contract that agrees to pay you one dollar for every dollar that stock A drops below $90 (think of the difference between $100 and $90 as the insurance deductible). If tomorrow the stock price drops from $100 to $95, that $90 insurance contract just became more valuable, and you can be sure its premium will rise in the market. Over time, the insurance premium fluctuates as news about the reference asset arrives. When the price of A drops, the insurance premium rises. In this case, you can replicate the insurance contract by taking a time-varying short position in stock A.

How does all of this relate to climate hedging? Suppose I can invest in two stocks, call one "brown" and the other "green." In the absence of a climate disaster, these stocks are identical, they both rise and fall with a beta of 1 to the market. However, in the event of a climate disaster, brown's value drops by half and green's value doubles. These would be great assets to be able to trade if I am interested in building a climate hedge portfolio! I would short $1 of brown and buy $1 of green. Absent a climate disaster, I have a zero-risk portfolio because the portfolio is neutral when there is no climate disaster. Should a climate disaster strike, brown drops, green rises, so both the long and short sides of my portfolio make money.[1] We have replicated climate disaster insurance, and did it simply by trading stocks and without requiring an explicit climate insurance market. Furthermore, (and this is where the replication argument in fact becomes relevant), my hedge portfolio will change in value not just when a climate disaster occurs, but any time any information arrives about the likelihood of a future disaster. Suppose tomorrow new information arrives that increases the probability of a disaster over the next year from say 1% to 2%. While the disaster has not occurred, and in this example still has a low chance of occurring, my portfolio will still appreciate tomorrow. A higher probability of disaster makes green that much more valuable, brown that much less valuable, and my hedge pays off already.

This is clear enough in theory. But how do I operationalize such a portfolio? If I want to establish a hedge for climate risk, how can I know which assets pay off when bad news about the climate arrives? This is where statistics comes in. The insight of Engle et al. (2020) is that

1 The ability to short is an important piece of the hedging puzzle. If we were to simply go long green stocks the portfolio would be dominated by market risk, rather than the climate risk sensitivity we are trying to isolate. In addition, we'd only pick up the upside sensitivity of the green stocks when a climate shock occurs, rather than also benefiting from the downside sensitivity of brown stocks by holding short positions. See also Palazzolo, Pomorski, and Zhao (2021).

we can establish effective climate hedges without relying on explicit insurance against future disasters. Rather, we can rely on the aggregate wisdom of the market to instruct us where – and how much – climate risk and opportunity resides. We can construct portfolios by taking long positions in securities that appreciate when we get bad news about climate change, while shorting assets that depreciate. We can do this even with short-term news arrival. As long as relevant news is arriving, we can use price responses in the market to identify stocks that are positive and negative climate change hedges.

Identifying Climate News

In order to identify which assets are effective for constructing a hedge portfolio, we need to know which assets respond positively and negatively to climate change news. A useful framework for thinking about the climate-hedging benefits of individual assets from a statistical standpoint is

$$r_{i,t} = \beta^{CC}_{i,t-1} CC_t + \beta_{i,t-1} F_t + \epsilon_{i,t} \qquad (1)$$

Equation 1 describes the behavior of a return on stock i in month t. It will be determined in large part by the stock's exposure to non-climate systematic factors, F_t (including market, size, and value factors, for example). The core of the model is stock i's sensitivity to negative climate change news, where the arrival of negative climate change news is represented by a new factor, CC_t (the more negative climate news arrives the higher is CC_t). The goal is to estimate the (conditional) betas of each stock on negative climate change news, $\beta^{CC}_{i,t-1}$, as these betas dictate the asset's role in a hedge portfolio. Assets with high values of $\beta^{CC}_{i,t-1}$ pay off when bad news arrives and are thus assigned proportionally large positive weights in the climate-hedge portfolios. Assets with negative $\beta^{CC}_{i,t-1}$ become short positions in the hedge portfolio.

The viability of a hedge portfolio hinges critically on the factor that measures the arrival of negative climate change news, CC_t. As Engle et al. (2020) note, "How should we identify the news sources that reflect the information investors use in their climate risk-based investment decisions? Once we identify the appropriate news, how do we measure its relative intensity over time? How do we quantify the extent of good news versus bad news? And should one differentiate among subtypes of climate news (such as news about physical climate risks versus news about regulatory risks)?"

A central contribution of Engle et al. (2020) is their proposal for CC_t. They propose two indices, both based on the intensity of climate change coverage in major print news media outlets.

The first index measures the coverage of climate change news by *The Wall Street Journal* (WSJ). The choice to build a climate change news index from WSJ is based on the comparatively high relevance of WSJ news for the investment community compared to other print media outlets.

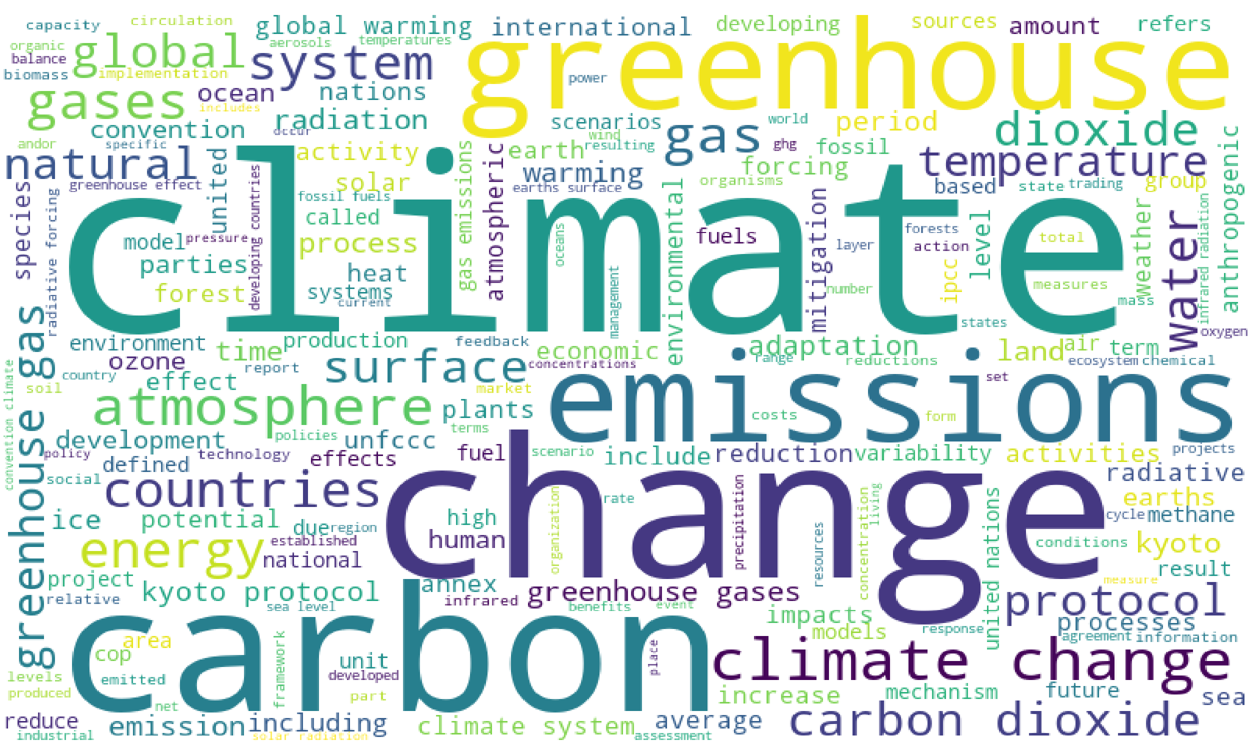
From an interpretation standpoint, the WSJ Climate Change News Index assumes that high coverage must be negative climate change news, since media attention is generally elevated when there is cause for concern, or when there is increased regulation – both of which would be expected to hurt brown stocks relative to green. Said differently, climate change is a theme with an inherent negative sign, so overall climate change coverage can be used as a proxy for negative climate change coverage.[2]

Another approach would be to identify the sentiment of news articles. Engle et al. (2020) study a second climate change news index extracted from a data analytics software called Crimson Hexagon. This index is designed to specifically focus on climate news bearing *negative* sentiment and is drawn from a wider array of news outlets. The results and conclusions of their analysis are similar for both WSJ and Crimson Hexagon indices. Thus, for brevity, we focus discussion on the WSJ index herein.

Both the WSJ and Crimson Hexagon climate change news indices proposed have their limitations. But, as emphasized by Engle et al. (2020), the power of this analysis lies in the generality of the method, as other innovative ideas for hedging targets can be easily analyzed and compared to those in Engle et al. (2020).

Constructing the WSJ Climate Change News Index requires the ability to quantify coverage of climate change, but how does one know what climate change news looks like? Engle et al. (2020) define climate change discourse based on word count frequencies in authoritative texts on the subject. This includes white papers from the Intergovernmental Panel on Climate Change and Environmental Protection Agency, as well as climate change glossaries from sources such as the United Nations. From these documents, the authors draw their "climate change vocabulary (CCV)" as a term cloud shown in **Figure 1**.

Figure 1: Climate Change Vocabulary



Source: Engle et al. (2020). Word cloud summary of climate change vocabulary from a corpus of seventy-four authoritative climate change texts. Term sizes are proportional to their frequency in the corpus.

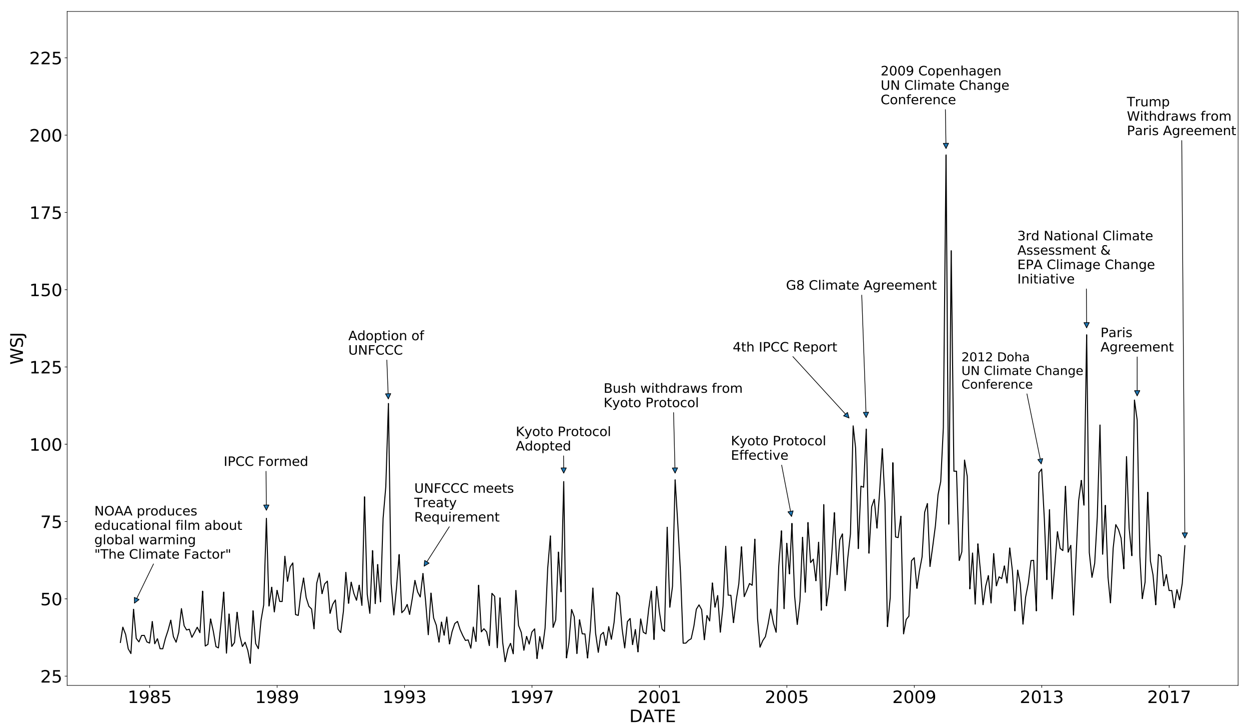
2 Of course not all climate change news is negative, and the sign of the impact on climate-sensitive stocks is not always clear. For example, news around the Paris Agreement is positive climate change news, but what is the impact on a climate hedge portfolio? The Paris Agreement might represent an increase in regulatory risk, which would increase the value of the hedge portfolio, but also a decrease in physical risk, which would have the opposite effect. Engle et al. (2020) note that distinguishing between different types of climate news is an interesting avenue for future research.

This summarizes the frequency distribution for terms (one-word and two-word phrases) related to climate change, with term sizes shown in proportion to their frequency of use.

With the climate change vocabulary pinned down, news coverage of climate change is measured as the similarity between CCV and term counts in *The Wall Street Journal* each day. More specifically, it is calculated as the cosine similarity[3] between "term frequency–inverse document frequency (tf-idf)" of CCV and WSJ.[4] If an edition of WSJ were to use terms in exact proportion with the climate change vocabulary, the index would have a value of one for that day, while a day with no overlap between WSJ and CCV would have an index value of zero. In other words, the index is designed to capture the fraction of the WSJ allocated to climate change discourse each day.

The resulting index, shown in **Figure 2**, describes the measure of climate change discourse over time, and innovations in this index constitute the hedging target (our estimate of CC_t).[5] It has intuitive time series behavior. Since its first availability in 1984 (the start of our full text WSJ sample), climate news coverage in WSJ has risen steadily over time.

Figure 2: WSJ Climate Change News Index



Source: Engle et al. (2020). This figure shows the level of the WSJ Climate Change News Index from 1984 to 2017 (scaled by a factor of 10,000), annotated with climate relevant news announcements. The level of the index is proportional to the fraction of WSJ coverage that is deemed to be climate related.

3 Cosine similarity is a mathematical method that can be used to measure the similarity between two bodies of text. For a description of this method see Appendix.
4 Common terms that appear in most documents earn low scores because they are less informative about any individual document's content (they have low idf), as do terms that are rare in a given article (they have low tf). The tf-idf transformation defines the most representative terms in a given document to be those that appear infrequently overall, but frequently in that specific document. For a discussion of the statistical advantages of tf-idf text representations, see Gentzkow, Kelly, and Taddy (2019).
5 This is similar to measuring changes in the index, but using innovations (specifically residuals from an AR(1) model) controls for mean reversion inherent in the index (if there's lots of climate change news one day, there's likely less climate change news the next day).

Moreover, it shows that climate news coverage intensifies around important climate events such as global climate summits (e.g., Kyoto, Copenhagen, and Paris) and other impactful policy events (e.g., Trump's withdrawal from the Paris Agreement).

Constructing the Hedge

Once the hedging target is determined there are multiple approaches one could take, under many different assumptions, to build a hedge portfolio. Herein we focus on the approach taken by Engle et al. (2020), which can be thought of as one reasonable framework for solving this problem that can be built upon in future research.

The analysis proceeds by defining CC_t in equation 1 (the hedging target) to be innovations in the WSJ Climate Change News Index and fixing F_t to be the overall U.S. market return. The authors use U.S. listed stocks as the base assets from which to build portfolios.

Conceptually, the key inputs to forming a hedge portfolio are stocks' climate news betas, $\beta^{CC}_{i,t\text{-}1}$ (their sensitivity to climate news). In the organizing framework of equation 1, these estimated betas are allowed to vary through time.[6] This flexibility accommodates changes in firms' business lines and climate policies, as well as evolution in the way climate risk impacts asset prices.

Engle et al. (2020) note that, to produce a hedging portfolio, the framework of equation 1 can be converted into a direct, one-step estimation problem that takes the form

$$CC_t = \xi + w_{t-1}{}'r_t + e_t \qquad (2)$$

where r_t is a vector of all individual stock returns traded at a point in time. The weights, w_t, are allowed to vary over time as a function of firm attributes, just like the dynamic betas in equation 1. When estimated via regression, the coefficients w_t are interpretable as the set of portfolio weights that best replicate—i.e., that represent the best hedge of—climate change news. The fitted value from this regression is our *climate news hedging portfolio*. Via regression equation 2, the weights w_t are constructed such that this portfolio has high returns when bad climate news arrives.

Naturally, the weight that this procedure places on stock i ($w_{i,t}$) has a close theoretical connection to the stock's climate news beta ($\beta^{CC}_{i,t\text{-}1}$) in equation 1. Stocks with high positive beta to climate news tend to appreciate when climate news (which is our proxy for increased climate risk) arrives, and stocks with strong negative beta will tend to drop in price on the same news. Thus, stocks with high absolute betas form successful hedges. Stocks with high positive sensitivity will have high values of $w_{i,t}$ and thus become overweights in the hedge portfolio, and stocks with large negative sensitivity will become the largest underweights.

Estimating time-varying weights or betas is an empirical challenge and prone to data mining.

6 A different way to tackle this problem would be to estimate each stock's climate news beta using a time series regression, similar to Taylor, Pastor, and Stambaugh (2021) for their 'green factor'. Under this approach the estimated betas, and therefore portfolio weights, are static when using the full sample. Engle et al. (2020) instead choose an approach that allows for dynamic weights, by estimating constant scalars which are applied to dynamic portfolios. This allows the hedge portfolio weights to reflect the dynamics of climate risk. One could also estimate dynamic betas using rolling samples, but we have a limited number of independent observations using monthly data so this would require higher frequency sampling.

The authors tackle this using the insights of Kelly, Pruitt, and Su (2019). In particular, betas are allowed to depend on a limited set of observable firm attributes that plausibly proxy firms' climate risk exposures, such as their environmental sustainability scores from ESG vendors such as MSCI (Z^{MSCI}) and Sustainalytics (Z^{SUS}). Three additional factors, value (Z^{HML}), size (Z^{SIZE}), and the market (Z^{MKT}), are also included in the regression since they might be correlated with climate risk and they are known for describing the cross-section of returns.[7] Taken altogether, the authors operationally implement equation 2 with the following time-series regression:

$$CC_t = \xi + w_{SUS} Z_{t-1}^{SUS\prime} r_t + w_{HML} Z_{t-1}^{HML\prime} r_t + w_{SIZE} Z_{t-1}^{SIZE\prime} r_t + w_{MKT} Z_{t-1}^{MKT\prime} r_t + e_t \quad (3)$$

where w_{SUS}, w_{HML}, w_{SIZE} and w_{MKT} are scalars that capture the weight of the corresponding portfolios in the hedge portfolio for CC_t. In short, we are solving for the weight of each characteristic-sorted portfolio that will best explain variation in climate news arrival. These estimated weights will then be our best guess of how to weight the four characteristic-sorted portfolios in equation 3 to form our hedge portfolio going forward.

Climate Risk Hedging Performance

Table 1: Solving for the weight of each characteristic-sorted portfolio in the hypothetical hedge portfolio; *three full-sample time-series regressions*

	Regression 1	Regression 2	Regression 3
$Z_{t-1}^{SUS\prime} r_t$	67.789***		
$Z_{t-1}^{MSCI\prime} r_t$		53.743*	
r_t^{XLE}			0.085
r_t^{PBD}			0.208
$Z_{t-1}^{HML\prime} r_t$	2.309	-5.941	-6.772
$Z_{t-1}^{SIZE\prime} r_t$	-6.034**	-5.459**	-2.765
$Z_{t-1}^{MKT\prime} r_t$	0.789	0.789	0.091
Constant	2.673	4.891*	5.959**
R^2	0.187	0.088	0.047
N	88	88	88

Source: Engle et al. (2020). This table shows results from regression equation 3 in column 1. Columns 2 and 3 use the same equation except $w_{SUS} Z_{t-1}^{SUS\prime}$ is replaced with $w_{MSCI} Z_{t-1}^{MSCI\prime}$ and $w_{XLE} Z_{t-1}^{XLE\prime} + w_{PBD} Z_{t-1}^{PBD\prime}$ respectively. The dependent variable captures innovations in the WSJ-Based Climate News measure. The unit of observation is a month, and the sample runs between September 2009 and December 2016. *p<.1; **p<.05; ***p<.01 (meaning that the estimate is statistically significant at the 10%, 5% and 1% significance level respectively). Estimated betas in column 1 are w_{SUS}, w_{HML}, w_{SIZE} and w_{MKT} and represent scalars that capture the weight of the corresponding portfolios in the hedge portfolio for CC_t. The magnitude of estimated betas represents the relative position size of each characteristic-sorted portfolio in the estimated hedge portfolio. A description of the universe of assets used and a description of the characteristic-sorted portfolio construction is provided in the Appendix. For illustrative purposes only and not representative of a portfolio AQR currently manages. Hypothetical data has certain inherent limitations, some of which are disclosed in the Appendix hereto.

7 We will see that size does in fact predict covariances between stocks and climate news.

Table 1 estimates equation 3 and summarizes the in-sample behavior of the WSJ-based climate change news hedge over the full sample 2009–2016. Column 1 shows that a hypothetical portfolio based on Sustainalytics E-Scores (Z^{SUS}) has a positive and significant sensitivity, and therefore ability, to hedge WSJ climate news. This indicates that in periods with more climate news, it is advantageous to be overweight "greener" firms with higher E-scores. Such a portfolio (combined with the other portfolios weighted based on this regression) can hedge roughly 18% of the in-sample variation in the WSJ climate news index (as measured by the regression R^2, which implies a correlation of 43% between CC_t and the hedge portfolio). Column 2 shows a similar result for a hedge based on MSCI E-Scores (Z^{MSCI}). In addition to the E-scores, the portfolio sorted on size (Z^{SIZE}) also correlates with CC_t, which implies that larger firms are more exposed to climate change news—they tend to realize lower returns than small firms around bad climate news events. Column 3 shows the association between CC_t and two popular ETFs that one would expect to have a close association with climate change news, the S&P 500 energy sector ETF (XLE) and the Invesco Global Clean Energy ETF (PBD).[8] These indices form a basic hurdle for our hedge portfolios, since one might expect that a long position in clean energy (PBD) coupled with a short position in traditional energy (XLE) would serve as a reasonable hedge to climate risk, and we see that each of the climate news hedge portfolios achieve a closer fit to CC_t than these two ETFs combined.

What would the hedge portfolios implied by Table 1 look like? The hedge portfolios are designed to be relatively industry-balanced, identifying those firms with the largest exposure to climate change risk both within and across industries. However, to provide some insight into positioning, we present the largest positive and negative average industry positions for the Sustainalytics regression specification (Table 1, column 1) in Table 2. The two largest shorts are "Coal Mining" and "Water Transportation," while the largest longs are "Building Materials and Gardening Supplies" and "Tobacco Products". These results highlight that the optimal climate change hedge portfolio isn't necessarily consistent with simple, intuitive rules of thumb, such as going long green energy stocks and shorting oil companies. The overweight to tobacco is particularly surprising since this industry is generally considered to be anti-ESG and is outright excluded in many 'sustainable' portfolios. It's important to note that the hedging portfolio is primarily tilted towards climate news sensitive stocks *within* industry, which provides valuable climate news sensitivity that is not visible in the table below.

8 The ETFs shown are used in the analysis by Engle et al. (2020) and do not form a recommendation by AQR. The results are for illustrative purposes only.

Table 2: Largest Average Hypothetical Short and Long Positions (By 2-Digit SIC Code)

Top Negative Portfolio Weights	SIC2 Code	Top Positive Portfolio Weights	SIC2 Code
Coal Mining	12	Building Materials & Gardening Supplies	52
Water Transportation	44	Tobacco Products	21
Insurance Agents, Brokers & Service	64	Food & Kindred Products	20
Mining Non-Metallic Minerals, Except Fuels	14	Paper & Allied Products	26
Transportation Services	47	Textile Mill Products	22

Source: Engle et al. (2020). This table shows the industries (2-digit SIC code) with the largest average short and long positions in the estimated hedge portfolio resulting from the regression presented in Table 1, column 1. Industries are arranged in ascending order of portfolio weights. The construction of the hedge portfolio is described herein and a description of the universe of assets used is provided in the Appendix. For illustrative purposes only and not representative of a portfolio AQR currently manages. Hypothetical data has certain inherent limitations, some of which are disclosed in the Appendix hereto.

The results of Table 1 are in-sample. In-sample analysis has the benefit of leveraging the longest possible time series for evaluating hedge performance. But the true test of a hedge portfolio is its ability to hedge out-of-sample climate news. To this end, Engle et al. (2020) cross-validate[9] performance of their hedges by estimating the model in one sub-sample, and testing the hedge portfolio on data that is not used for estimation. We report the implied correlations from this robustness check in Table 3. The WSJ climate news hedge portfolio that uses Sustainalytics E-Scores (H_{OOS}^{SUS}) substantially outperforms the alternative on an out-of-sample basis with a 30% correlation to innovations in the WSJ climate news index (CC_t).[10] These alternatives include the WSJ hedge portfolio based on MSCI E-Scores (H_{OOS}^{MSCI}) and the pre-existing ETFs (r_t^{XLE}) and (r_t^{PBD}).

Table 3: Out-of-sample correlations of hypothetical hedge portfolios with climate news

	H_{OOS}^{SUS}	H_{OOS}^{MSCI}	r_t^{XLE}	r_t^{PBD}
CC_t	0.300	0.067	0.068	0.111

Source: Engle et al. (2020). This table shows out-of-sample cross-validation correlation of different hedge portfolios and innovations in the WSJ Climate Change News Index. The construction of the hedge portfolios is described herein and a description of the universe of assets used is provided in the Appendix. For illustrative purposes only and not representative of a portfolio AQR currently manages. Hypothetical data has certain inherent limitations, some of which are disclosed in the Appendix hereto.

9 Cross-validation is a technique used to assess the effectiveness of a model when there is limited data available by 'training' the model on various sub-sets of the full-sample and 'testing' it on the remaining sample in each case, then taking an average of the errors. The technique is used to prevent overfitting or 'data-mining'.
10 As a reminder, the implied in-sample correlation from Table 1 column 1 was approximately 43%.

Conclusions

Engle et al. (2020) present a novel and effective climate-risk hedging portfolio derived from a combination of economic theory and textual analysis. While their emphasis is on hedging climate change news coverage in *The Wall Street Journal*, the framework is a general, rigorous methodology. It can flexibly accommodate alternative hedging targets that researchers might hypothesize. Indeed, Engle et al. (2020) emphasize that their framework is not a definitive climate-hedging solution, but is instead a launching point for exploring a climate-hedging agenda.

We stress that the framework is not meant as a replacement for carbon-aware investing. Climate risk is extremely uncertain and complex, so there's no way of modelling this risk precisely.[11] When we're faced with model uncertainty there are diversification benefits to incorporating information from multiple partially correlated models. Carbon emissions are one practical, but imperfect, model of climate risk as is the model presented herein: climate hedges based on correlation with climate news. We believe that combining multiple views of climate risk will play an increasingly important role in investors' climate hedging toolkit.

References

Engle, Giglio, Kelly, Lee, and Stroebel, 2020, "Hedging Climate Change News", working paper.

Palazzolo, Pomorski and Zhao, 2021, "(Car)Bon Voyage: The Road to Low Carbon Investment Portfolios", AQR whitepaper.

Nordhaus, 2014, "Estimates of the Social Cost of Carbon: Concepts and Results from the DICE-2013R Model and Alternative Approaches", *Journal of the Association of Environmental and Resource Economists*, Vol. 1, No. 1/2 (Spring/Summer 2014), pp. 273-312.

Sommervoll and Swidler, 2021, "Hedging Home Equity Risk: Examination of a Nobel Idea", *Journal of Housing Research*, DOI: 10.1080/10527001.2021.1985370.

Black and Scholes, 1973, "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, Vol. 81, No. 3 (May - Jun., 1973), pp. 637–654.

Merton, 1973, "Theory of Rational Option Pricing", *The Bell Journal of Economics and Management Science*, Vol. 4, No. 1 (Spring, 1973), pp. 141-183.

Gentzkow, Kelly, and Taddy, 2019, "Text as Data", *Journal of Economic Literature*, 57(3), pp. 535–574.

Taylor, Pastor, Stambaugh, 2021, "Dissecting Green Returns", working paper.

Giglio, Kelly, and Stroebel, 2020, "Climate Finance", working paper.

11 Giglio, Kelly, and Stroebel (2020).

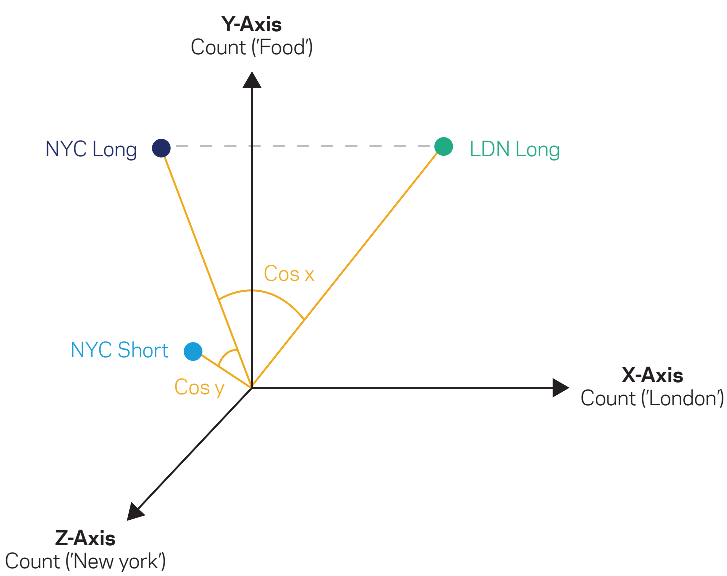
Appendix

Cosine Similarity Applied to Textual Analysis

Cosine similarity is a mathematical technique that can be used, in this case, to measure the similarity between corpora of text. First let's set up the problem, how can one systematically measure the similarity of two documents? One way would be to count the number of times words or phrases appear in both documents, but there's a problem, the longer the documents the more likely they are to have high common word counts. What we really want to know is the similarity of the ratio in which words appear in each document, and we can measure that using cosine similarity (in math speak this is the cosine of the angle between vectors where the vectors are arrays of word counts for each corpus of text).

Here's an example: assume we have three texts on the best places to eat in particular cities; the first is a magazine on the best places to eat in London, the second is the same magazine for New York, and the third is an online advertisement for the New York magazine. To measure the similarity between the texts we could count the number of times that the words 'London', 'New York' and 'Food' appear in each, and plot them on three axes as shown in **Figure A**. Simply counting the number of common words in each text would measure the distance between points[12] (such as the dotted grey line), but as we said earlier that could lead us to misinterpret longer documents as being more similar than they actually are (e.g. the London and New York magazines could be more similar on this measure than the New York magazine and its advertisement). So instead we measure the cosine of the angles between vectors which are shown by *Cos x* and *Cos y* below. Now we can clearly see that the texts about the best places to eat in New York are more similar than the text about London.

Figure A: Illustration of Word Counts in 3-Dimensional Space



We can then extend this simple illustration to include many more words, extending the problem across multiple dimensions. This is more difficult to visualize but the same math applies, and Engle et al. (2020) lean on this framework to identify the similarity between *Wall Street Journal* articles and their Climate Change Vocabulary when building the WSJ Climate Change News Index.

Source: AQR. NYC Long represents the New York magazine, NYC Short represents the New York magazine advertisement, LDN Long represents the London magazine. Each point is plotted based on the word count of 'London', 'Food', 'New York'. For illustrative purposes only.

12 Formally known as the Euclidean distance approach.

Disclosures

Data description:

Engle et al. (2020) focus on constructing hedge portfolios using U.S. equities as the underlying assets. They obtain monthly individual U.S. stock return data from CRSP and include only common equity securities (share codes 10 and 11) for firms traded on the NYSE, AMEX and NASDAQ. They exclude penny stocks, defined as stocks with a price below $5 at the time of portfolio formation. This is to avoid including stocks whose returns are dominated by market microstructure issues. They also drop microcap stocks, defined as stocks with market capitalization in the bottom 20% of the sample traded in NYSE.

MSCI E-scores are from a data set of annual firm-level ESG scores between 1995 and 2016 from MSCI. To convert to monthly scores the authors assign the same score to all the months in the relevant year. Sustainalytics E-scores are monthly firm-level ESG scores beginning in September 2009 from Sustainalytics.

Portfolio descriptions:

Size portfolio is constructed using cross-sectionally standardized market value so that half the firms, sorted by market value, have positive weight, and half have negative weight; note that this portfolio will be long large firms and short small firms). Value portfolio is constructed using cross-sectionally standardized values of book-to-market. Market portfolio weights are equal to the share of total market value. E-score portfolios are constructed by ranking the E-scores of all firms at each point in time, and then demeaning and re-scaling the ranked measures such that they range from -0.5 to +0.5. The authors also present the results of a cross-sectional demeaning of each E-Score in each month in their paper.

Fees or transaction costs are not taken into consideration in the analysis herein.

This presentation is not research and should not be treated as research. This presentation does not represent valuation judgments with respect to any financial instrument, issuer, security, or sector that may be described or referenced herein and does not represent a formal or official view of AQR.

The views expressed reflect the current views as of the date hereof, and neither the author nor AQR undertakes to advise you of any changes in the views expressed herein. It should not be assumed that the author or AQR will make investment recommendations in the future that are consistent with the views expressed herein, or use any or all of the techniques or methods of analysis described herein in managing client accounts. AQR and its affiliates may have positions (long or short) or engage in securities transactions that are not consistent with the information and views expressed in this presentation.

The information contained herein is only as current as of the date indicated and may be superseded by subsequent market events or for other reasons. Charts and graphs provided herein are for illustrative purposes only. The information in this presentation has been developed internally and/or obtained from sources believed to be reliable; however, neither AQR nor the author guarantees the accuracy, adequacy, or completeness of such information. Nothing contained herein constitutes investment, legal, tax, or other advice, nor is it to be relied on in making an investment or other decision.

There can be no assurance that an investment strategy will be successful. Historic market trends are not reliable indicators of actual future market behavior or future performance of any particular investment, which may differ materially, and should not be relied upon as such. Target allocations contained herein are subject to change. There is no assurance that the target allocations will be achieved, and actual allocations may be significantly different from those shown here. This presentation should not be viewed as a current or past recommendation or a solicitation of an offer to buy or sell any securities or to adopt any investment strategy.

The information in this presentation might contain projections or other forward-looking statements regarding future events, targets, forecasts, or expectations regarding the strategies described herein and is only current as of the date indicated. There is no assurance that such events or targets will be achieved and might be significantly different from that shown here. The information in this presentation, including statements concerning financial market trends, is based on current market conditions, which will fluctuate and may be superseded by subsequent market events or for other reasons. Performance of all cited indices is calculated on a total return basis with dividends reinvested.

The investment strategy and themes discussed herein may be unsuitable for investors depending on their specific investment objectives and financial situation. Please note that changes in the rate of exchange of a currency might affect the value, price, or income of an investment adversely. Neither AQR nor the author assumes any duty to, nor undertakes to update forward-looking statements. No representation or warranty, express or implied, is made or given by or on behalf of AQR, the author, or any other person as to the accuracy and completeness or fairness of the information contained in this presentation, and no responsibility or liability is accepted for any such information. By accepting this presentation in its entirety, the recipient acknowledges its understanding and acceptance of the foregoing statement. Diversification does not eliminate the risk of experiencing investment losses.

AQR

www.aqr.com

AQR Capital Management, LLC Two Greenwich Plaza, Greenwich, CT 06830 **P** +1.203.742.3600 **F** +1.203.742.3100