Does Environmental Policy Uncertainty Hinder Investments Towards a Low-Carbon Economy?

Joëlle Noailly[a,b,c*] Laura Nowzohour[a]

Matthias van den Heuvel[d]

[a]Geneva Graduate Institute, Switzerland
[b]Vrije Universiteit Amsterdam, The Netherlands
[c]Tinbergen Institute, The Netherlands
[d]EPFL, Switzerland

July 2022

Abstract

We use machine learning algorithms to construct a novel news-based index of US environmental and climate policy uncertainty (EnvPU) available on a monthly basis over the 1990-2019 period. We find that the EnvPU index spikes during the environmental spending disputes of the 1995-1996 government shutdown, in the early 2010s due the failure of the national cap-and-trade climate bill and during the Trump presidency. We examine how elevated levels of environmental policy uncertainty relate to investments in the low-carbon economy. In firm-level estimations, we find that a rise in the EnvPU index is associated with a reduced probability for cleantech startups to receive venture capital (VC) funding. In financial markets, a rise in our EnvPU index is associated with higher stock volatility for firms with above-average green revenue shares. At the macro level, shocks in our index lead to declines in the number of cleantech VC deals and higher volatility of the main benchmark clean energy exchange-traded fund. Overall, our results are consistent with the notion that policy uncertainty has adverse effects on investments for the low-carbon economy.

JEL-Classification: Q58; C55; D81; E22

*Corresponding author: Joëlle Noailly, joelle.noailly@graduateinstitute.ch, Tel: +41229086222, Address: CIES, Geneva Graduate Institute, Chemin Eugene-Rigot 2, 1202 Geneva, Switzerland. This study is supported by the Swiss National Science Foundation (SNSF) within the framework of the National Research Programme "Sustainable Economy: resource-friendly, future-oriented, innovative" (NRP 73, Grant-No 407340-172395). The research also benefited from a subaward granted via the 2020 NBER Workshop on Economics of Innovation in the Energy Sector supported by the Sloan Foundation. We thank Patrick Ruch, Luc Mottin, Julien Gobeill (Swiss Institute of Bioinformatics) for comments and research assistance on text-mining algorithms. In addition, we thank F. Duriaux, K. Purevsuren, J. Suda, V. Albertolli, C. Earle, R. Taniguchi for excellent research assistance, as well as participants at the 2020-2021 NBER Economics of Energy Innovation workshops, Vrije Universiteit Amsterdam and Coventry University seminars, and the 2022 Morzine NRP73 Research Retreat, IAERE, EAERE, and EEA meetings for helpful comments.

1 Introduction

Over the course of its mandate, the Trump administration not only announced its withdrawal from the Paris climate agreement but also conducted a comprehensive review of many federal environmental regulations. These changes created a surge in uncertainty about the state of future environmental and climate regulations as it became unclear how and when they would be implemented, how many policies would be dismantled and whether these rollbacks would be legally challenged in the future. Such policy turnarounds are not unique to the US. France revoked its fuel tax after the Yellow Vests protests and Australia and China have also backpedalled on coal regulations in recent years. While policy revisions in response to new information are inevitable and desirable, the politics of environmental and climate policy are particularly volatile. Environmental regulations typically face a lot of opposition in the form of lobbying and protests and policymakers often have to trade long-term environmental objectives for short-term economic and electoral priorities. The history of environmental and climate regulations shows many episodes where even advanced and promising policy proposals have in the end failed unexpectedly.

Yet, such abrupt policy changes generate substantial uncertainty, making it difficult to anticipate how the regulatory framework will unfold in the future. Faced with high levels of uncertainty about future environmental and climate policy, firms and investors may prefer to adopt a *wait-and-see* behavior and refrain from investing in the low-carbon economy – in particular, as these investments strongly rely on public policies to be profitable. According to a recent survey from the European Investment Bank, 43 percent of European firms and 22 percent of US firms cite 'uncertainty about regulation' as an important barrier to undertaking climate-related investment. European firms rank policy uncertainty as the most important obstacle, while for US firms policy uncertainty comes just after 'investment costs' but above 'availability of finance' (European

Investment Bank, 2021). Given the urgency of climate action, understanding how policy uncertainty affects and delays clean investments is essential to provide better guidance on the significance of the timing and credibility of environmental and climate regulations (Goulder, 2020).

The objective of this paper is to introduce a novel news-based index of US environmental and climate policy uncertainty at the monthly level and to examine the adverse effects of such policy uncertainty on investments for the low-carbon economy. We first review the definition, measurement, accuracy and validation of our index. We define environmental and climate policy uncertainty as the inability to predict how the regulatory regime on environmental and climate issues will unfold in the future. We consider a broad scope of environmental and climate policies entangling both state and federal regulations on a wide range of environmental concerns.[1] We measure environmental and climate policy uncertainty by using supervised machine learning algorithms on the text of news articles extracted from the archive of ten US newspapers over the 1990-2019 period. Readings of news articles show that many factors give rise to perceptions of environmental and climate policy uncertainty in the US context, such as the unpredictable outcome of the legislative process, legal challenges of regulations awaiting for court decisions, unexpected revision or rollback of policies or the failure of international environmental negotiations. Accordingly, we train the algorithm with manually coded news articles reflecting perceptions of elevated levels of uncertainty about current and future policy. We exclude articles referring to past or declining uncertainty, as well as news on other forms of non-policy related uncertainty (such as uncertainty related to climate change impacts). Using the set of articles predicted by our best performing classifying algorithm, we compute the EnvPU index as the monthly frequency of policy uncertainty articles in the volume of environmental policy news.

[1]We consider regulations of environmental pollution (e.g. greenhouse gas emissions and other air pollutants from electricity generation, vehicles and buildings, water pollution, oil spills, toxic and hazardous waste) and abstracts from policies regulating natural resources (e.g. forests, fishery, groundwater extraction).

Our methodology based on supervised machine learning techniques addresses concerns about the accuracy of the index. In particular, we find that our method outperforms other dictionary and keyword-based methods as our classifying algorithm is better able to capture variations in semantics and topics related to policy uncertainty. We then present a series of validity checks. We first show that spikes in our index corresponds to well-known historical episodes. Our EnvPU index spikes at the end of 1995 when a disagreement over cuts in environmental regulations led to a government shutdown for several weeks. The early 2010s are also punctuated by several bursts of policy uncertainty related to legislative hurdles around the climate bill aiming to introduce a national cap-and-trade system, the failure of the COP15 in Copenhagen, legal challenges between Texas and the EPA, and new rules regarding offshore drilling after the Deepwater Horizon oil spill. Finally, the EnvPU index surges at record-high levels during the Trump presidency. Next, we conduct a human audit to corroborate that our index is correlated with manual labeling by external auditors. We also verify that our index is not affected by political slant in newspaper coverage. Finally, we document in more detail the relationship between our EnvPU index and US election cycles, since elections are major events causing unpredictability in the regulatory framework.

Finally, we examine whether environmental policy uncertainty hinders investments in the low-carbon economy. Using firm-level data, we find that higher values of the EnvPU index are associated with (i) lower probability of receiving venture capital (VC) financing for most exposed startups active in cleantech sectors, (ii) higher volatility of stock returns for firms investing in the low-carbon economy. Using vector autoregressive (VAR) models, we find similarly that a shock in EnvPU leads to reduced clean energy VC deals at the aggregate level and higher volatility of the main clean energy exchange-traded fund. These results suggest that environmental policy uncertainty has adverse effects on investments for the low-carbon economy.

Our study presents a novel news-based index of environmental and climate policy uncertainty using machine learning techniques. The EnvPU index is freely accessible online at www.financingcleantech.com/envpu-index. Our work connects to the growing number of studies relating news-based indices to economic outcomes. While several of these indices measure policy uncertainty in other domains – economic policy uncertainty (Baker et al., 2020), trade policy uncertainty (Caldara et al., 2020), and geopolitical risks (Caldara and Iacoviello, 2022) among others – the focus on environmental and climate regulations is relatively new. To our knowledge, only two very recent studies have searched specifically for climate policy uncertainty articles in newspapers (Gavriilidis, 2021; Basaglia et al., 2021).[2] We differ from their work by using machine learning techniques rather than dictionary-based approaches on a broader set of newspapers and by considering a wider range of environmental regulations beyond climate change (including other air pollutants beyond greenhouse gas emissions, water pollution, oil spills, toxic and hazardous waste, etc). Our index provides thus a more accurate and richer measure of environmental and climate policy uncertainty. We are also the first paper to show the role of elections and shifts in partisan balance in US politics as an underlying channel of environmental and climate policy uncertainty.

Our study relates to the scant empirical literature in environmental economics exploring the effects of environmental policy uncertainty on financial investments around the timing of specific policy episodes. Lemoine (2017) exploits the unexpected collapse of the cap-and-trade climate bill in April 2010 to show that it led to an increase in coal prices and inventories. Sen and von Schickfus (2020) find that uncertainty about the implementation of a compensation mechanism for a carbon fee for energy companies led to an abrupt devaluation of companies holding fossil fuel assets. Dorsey (2019) examines the effects of policy uncertainty created by legal challenges to the Clean Air Interstate Rule and finds that plants with a lower probability of being regulated reduced pollution by 13 percent less and compliance costs overall increased by $124 million due to efficient

[2]Two additional studies by Engle et al. (2020) and our own previous work in Noailly et al. (2021) also extract information from newspapers on climate change and environmental policy news, respectively.

investments being delayed. While most of this work looks at single policies mostly using event studies, we add to this literature by bringing in new high-frequency data to track policy uncertainty over the long history of US environmental and climate policy. Such improved quantification also helps to shed light on the variety of circumstances giving rise to policy uncertainty (partisan disputes, legal challenges, elections).

Finally, our study contributes to the recent literature on climate change and finance by providing an improved quantification of regulatory risks associated to environmental and climate policies. There is a rapidly growing literature using text-as-data methods in economics (Gentzkow et al., 2017; Dugoua et al., 2022) and a few recent applications in finance aim to identify climate risks or more specifically transition risks by conducting textual analysis on 10k filings and earnings calls (Sautner et al., 2020; Kölbel et al., 2020). Transition risks refer to how firms, in particular fossil-fuel or highly-polluting firms, may be affected by the risks of a significant strengthening of environmental and climate policy in the future leading to the assets of those firms to be stranded. Our EnvPU index measures a different type of regulatory risk, namely the volatility and uncertain trajectory of future environmental and climate policy, without making specific projections on the direction of change in policy stringency. Our novel measure can thus improve the quantification of policy risks affecting financial portfolios.

The paper is organized as follows. Section 2 describes our methodology and text-mining algorithm and depicts the EnvPU index. Section 3 provides a set of validity checks. Section 4 explores how our index relates to low-carbon investments and Section 5 concludes.

2 Measuring Environmental and Climate Policy Uncertainty

In this section, we present the various steps towards constructing our news-based EnvPU index of environmental and climate policy uncertainty using automated methods. Our methodology is based on supervised machine learning techniques along a two-step approach. The first step consists in building a supervised learning algorithm able to identify the subset of environmental and climate policy news articles into the total volume of news. This stage is presented at length in earlier work (Noailly et al., 2021) and we provide a brief summary in Section 2.1. The next step that is the core of the present analysis proceeds in a similar fashion to construct a second supervised machine learning algorithm to classify policy uncertainty news within the subset of environmental and climate policy news. In this way, we are able to characterize both the level of environmental and climate policy news (first-moment) and the level of uncertainty in environmental and climate policy news (second-moment). We discuss the advantages and limitations of our two-step machine learning approach in Section 2.2.

2.1 Environmental and climate policy news

Our initial dataset is a sample of 80,045 newspaper articles about US environmental and climate policy, which we have identified in previous work (Noailly et al., 2021). The classification exercise started from 15 million news articles extracted from the archives of ten leading US newspapers over the period from January 1981 to March 2019 obtained via automated access through Dow Jones Factiva's platform. The list of newspapers includes: *New York Times, Washington Post, Wall Street Journal, Houston Chronicle, Dallas Morning News, San Francisco Chronicle, Boston Herald, Tampa Bay Times, San Jose Mercury News and San Diego Union Tribune.*[3]

[3]See Appendix A and Noailly et al. (2021) for descriptives and statistics on the set of newspapers, as well as on the method used to restrict the sample of articles.

The 80,045 news articles on environmental and climate policy have been identified by using a supervised support vector machine (SVM) algorithm, trained on a set of about 2,500 manually labeled articles. The decision rule applied by the trained algorithm out-of-sample identifies a set of 80,045 relevant articles. We scale the monthly counts of environmental and climate policy articles by the total monthly volume of news articles in our ten newspapers to construct an index of environmental and climate policy (the EnvP index). An in-depth presentation and validation of this index and how it positively correlates to policy stringency and low-carbon investments can be found in Noailly et al. (2021). We also show in previous work that the news articles underlying the EnvP index contains very rich and detailed information on a variety of policy topics, such as renewable energy, automobile emissions, water pollution, waste and recycling, green buildings, etc – much beyond climate change issues.

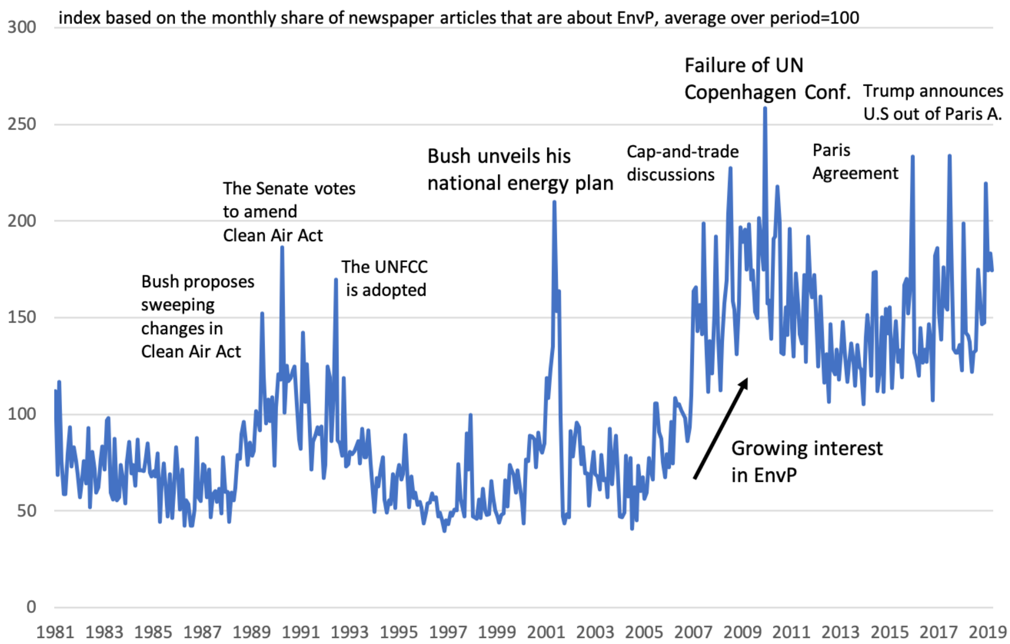
Figure 1 reproduces our index of environmental policy, which corresponds to the monthly share of all news articles that are about environmental policy, normalized to an average value of 100 over the 1981-2019 period. The index correctly captures salient events in the history of US environmental policy, such as the Green New Deal during Obama's Presidency, major UNFCCC climate change conferences, or Trump's announcement of withdrawal from the Paris agreement in June 2017. In the remainder of our analysis, we use the EnvP index to control for the level (first-moment) of environmental policy, whereas our newly constructed environmental policy uncertainty index reflects how volatile and unpredictable (second-moment) the regulatory framework is.

2.2 Developing an index of environmental policy uncertainty

Training set, text pre-processing and classification

Within the set of 80,045 EnvP articles described in Section 2.1, we aim to identify the subset of articles about environmental policy *uncertainty*. The first step in supervised learning methods is to build a manually annotated training set to inform the algorithm

Figure 1: EnvP - An index of environmental policy



about the content of relevant articles. We start by reading a large set of articles to develop a detailed codebook to guide the classification of EnvPU articles.

Typically, articles giving rise to perceptions of an inability to predict how future environmental policy will unfold fall into the following categories:

1. There is a major policy shift or reversal of environmental regulations, creating volatility in the politics of environmental policy. This type of uncertainty is inherent to elections and transitions in political cycles.[4]

2. There are various challenges threatening whether a given environmental regulation will be adopted and implemented. These challenges typically have to do with:

    - political or business opposition which may block or slow down the legislative

[4]Baker et al. (2020) show that national election cycles influence economic policy uncertainty, as measured by their EPU index. They show that increases in economic policy uncertainty are especially pronounced before close and highly polarised elections.

process and thereby the fate of the policy,

- a lack of political will and supportive coalitions, for instance in the case of climate change negotiations,
- legal challenges where the court's decision is still pending.

3. There is uncertainty about what the policy rules will exactly entail, when the policy will start, or whether the policy will be enforced.
4. Other types of policies (e.g. trade dispute on solar tariffs) may have uncertain impacts on clean markets.

We only label articles as relevant if they refer to rising levels of current and future environmental and climate policy uncertainty. Hence, we do not consider news articles which refer to past, declining or resolved policy uncertainty. In addition, we exclude articles referring to other forms of non-policy uncertainty – for instance, articles mentioning the uncertain impacts of climate change.

Guided by this codebook, we manually label a random sample of 622 articles from our subgroup of 80,045 EnvP news articles. Each article is labelled separately by at least two annotators. The result of the manual classification yields 204 articles labeled as relevant for environmental policy uncertainty, so about 30 percent of the training set. Our training set is relatively small and we may be worried that the sample may not be large enough to capture sufficient information on policy uncertainty. Nonetheless, we found that increasing the sample by additional increments did not significantly improve the performance of our classifying algorithm. Given that manually labeling articles is a very time-consuming and difficult task, we did not consider further extension of the training set at this stage.

In a second step, we apply standard text pre-processing techniques to our set of environmental policy news articles and convert each document into numerical vectors of unigram, bigram and trigram frequencies.[5] We then construct a term-frequency inverse-document frequency (tf-idf) matrix, in which less weight is given to words that occur too often or too rarely.[6]

Next, we input our preprocessed training set into a support vector machine (SVM), which is a supervised learning algorithm often used for text classification. The algorithm learns from the training set which text features are most (and least) important in determining whether an article pertains to environmental policy uncertainty or not.

Finally, we apply the prediction rule of our SVM classifier to the entire sample of 80,045 EnvP articles. This provides us with a new refined corpus of 25,174 newspaper articles on environmental policy uncertainty. Hence, around 31 percent of our EnvP articles are labelled as EnvPU, which is in line with insights from our manual labeling exercise.

Using our best performing algorithm, we obtain an average precision of 56 percent and a recall of 70 percent, when predicting which of the 622 articles in our training set are about EnvPU.[7] In other words, our precision metrics tells us that more than half of the articles classified as EnvPU were also labelled as uncertain by the annotators. Moreover, our recall metrics tells us that the EnvPU classifier successfully retrieve more than two thirds of all the articles labelled as uncertain.

How should we evaluate these metrics? First, with a precision of 56 percent, our algorithm performs significantly better than a random classifier (which would give a precision of about 30 percent given that less than one third of articles are classified as relevant). Second, although a precision of 56 percent might seem low at first sight, this

[5]We favor a bag-of-words approach over more sophisticated models relying on word vectors (word embedding), as it allows for more transparency into the decision rule of the classifier, allowing for easier interpretation and validation.

[6]See Noailly et al. (2021) for more details on pre-processing and matrix transformation of news articles.

[7]These performance metrics are an average of five different ten-fold cross validations using different random seeds. We use similar parametrization and cross-validation as in Noailly et al. (2021). Precision is the fraction of documents identified as relevant by the classifier that were indeed labeled as relevant by the annotators. Recall is the fraction of the relevant documents that are successfully retrieved by the classifier. The F1-score is 62 percent with a precision of 50 percent and a recall of 70 percent.

needs to be put in perspective with our two-step supervised learning approach. Our precision metrics is affected by the fact that the task of distinguishing environmental policy uncertainty articles is harder within a subset of environmental policy news than within a general set of news (the one-step approach commonly used in the literature). To illustrate this point further, it would be for instance much easier to obtain a precision of above 90% if we would try to identify an article on an unexpected reversal of climate legislation within a set of articles mixing foreign policy, sport and entertainment news. Yet, our approach is more arduous because we want to be able to retrieve these specific articles within the subset of relatively homogeneous news on environmental and climate policy. Inference processes are thus computationally harder in this case. Besides its practical benefit of limiting the number of articles to read (given that environmental policy uncertainty news are very rare in the total volume of news), we believe that our two-step approach is more accurate and less prone to errors than reading of (scarce) environmental policy uncertainty articles among all types of news. In our case, our annotators are better trained to understanding and distinguishing policy uncertainty in generic environmental policy discussions. Overall, identifying uncertainty is a complex and subjective task, which is difficult to categorize with a limited list of keywords. Our exercise shows that even with a codebook, two annotators would on average disagree on the label assigned to an article about 30 percent of the time. Hence, if humans cannot perfectly identify relevant articles, we cannot expect our algorithm to do so.

Descriptive statistics

Table 1 displays some of the most important text features used by the SVM classifier to predict whether an article falls into the 'environmental policy uncertainty' classification. As shown, this list of words encompasses 1) environmental and climate issues (i.e. 'emission', 'pipeline', 'drilling', 'auto', 'clean', 'wind'), 2) policy-making (i.e. 'epa', 'agency', 'rule', 'congress', 'administration', 'trump', 'clinton') and 3) uncertainty terms

Table 1: Top discriminating words for predicting our EnvPU index according to the trained SVM classifier.

Word	Weight	Word	Weight	Word	Weight
epa	1.77	cut	0.74	treaty	0.64
agency	1.24	trump	0.73	delay	0.64
rule	1.06	court	0.73	oil	0.64
state	0.93	new	0.71	regulation	0.64
congress	0.93	bill	0.69	economy	0.63
could	0.91	emission	0.69	canada	0.62
administration	0.91	clean	0.69	official	0.62
pipeline	0.90	wind	0.68	fracture	0.62
review	0.90	arpae	0.67	sand	0.61
permit	0.88	issue	0.67	federal	0.61
group	0.86	fight	0.67	lease	0.60
proposal	0.85	clinton	0.67	republican	0.60
drilling	0.81	acid	0.66	lead	0.60
law	0.78	txi	0.66	ballot	0.59
auto	0.75	forest	0.64	cape wind	0.58

(i.e. ‘court’, ‘review’, ‘cut’, ‘issue’, ‘fight’ or ‘delay’).

Table 2 reports excerpts of the five newspaper articles that were classified as the most likely to be about EnvPU. The first article titled *Trump officials deploy court tactic to reverse Obama rules* describes how President Trump was using both executives orders and court tactics to nullify numerous Obama-era green regulations. The second article *Court rebuffs Trump's effort to halt Obama methane rule*, reports on a federal court decision to prevent the EPA from suspending methane regulations. All of these articles describe either the rollback of environmental policies or legal battles over environmental regulations.

Table 2: Newspapers articles with the highest SVM-score

Title	Date	Score	Newspaper	Excerpt
Trump officials deploy court tactic to reverse Obama rules	Apr 18, 2017	2.14	Washington Post	”[...] President Trump has signed executive orders with great fanfare and breathed life into a once-obscure law to nullify numerous Obama-era regulations. But his administration is also using a third tactic: Going to court to stop federal judges from ruling on a broad array of regulations that are being challenged [...]”
Court Rebuffs Trump’s Effort To Halt Obama Methane Rule	Jul 4, 2017	2.07	New York Times	”[...] a federal appeals court ruled on Monday that the EPA cannot suspend an Obama-era rule to restrict methane emissions [...] The ruling signals that the Trump administration’s efforts to simply delay environmental and public health actions are likely to face an uphill battle in the courts and require a more painstaking process”
Rule-Making Process Could Soften Clean Air Act	Sep 21, 1991	2.07	Washington Post	”As the Senate neared passage of the new Clean Air Act last year, the Bush administration was pushing hard for inclusion of a special provision easing expensive pollution control requirements for electric utilities. [...] Administration efforts for the provision were rebuffed three times [...]”
23 Environmental Rules Rolled Back in Trump’s First 100 Days	May 3, 2017	1.93	New York Times	”President Trump, with help from his administration and Republicans in Congress, has reversed course on nearly two dozen environmental rules, regulations and other Obama-era policies during his first 100 days in office.”
Texas leads climate rules attack	Jan 11, 2011	1.86	Dallas Morning News	”Texas has filed nearly a dozen legal challenges of EPA regulations over the past year, mostly over climate-change rules. [...] Environmental groups say they expected that some states and business groups would continue to fight carbon limits, even after the Supreme Court’s decision [...]”

Next, we generate our index of US environmental policy uncertainty – the EnvPU index, which represents the monthly share of environmental policy uncertainty articles over all environmental and climate policy articles, normalized such that its average value over the 1990-2019 period is equal to 100. Hence, a rise in our EnvPU index captures a meaningful increase in the prevalence of policy uncertainty amid ongoing debates on environmental and climate policy. Our index is, therefore, unaffected by increasing media attention to environmental policy news. Scaling by the count of environmental policy articles improves over scaling by the total volume of news as the latter would not allow us to differentiate increases in EnvPU due to more coverage to environmental policy (which by construction would raise the count of articles about policy uncertainty) from increases due to an actual rise in policy uncertainty.

Figure 2 plots the historical evolution of US environmental policy uncertainty over the 1990-2019 period. We find that our EnvPU index peaks at the end of 1995 when a disagreement over cuts in environmental regulations led to a government shutdown for several weeks. In this period, Republicans gained control of both houses of Congress for the first time since 1954 and attempted to push an anti-regulations agenda and block Federal agencies from imposing new rules on health, safety and the environment. Environmental policy uncertainty falls in the mid-2000s, while the end of the decade is punctuated by several bursts in policy uncertainty.

Figure 3 zooms in on the 2009-2019 period to give a more fine-grained picture of what our index is able to capture. The first spike corresponds to the summer of 2010, during which our index was 40 percent above its average level. Uncertainty was then due to the introduction and eventual failure of a comprehensive climate bill sponsored by John Kerry, Joe Lieberman and, initially, Lindsay Graham. The second spike corresponds to the early months of Trump's presidency in 2017, where our index was 60 percent above its average level. At the time, there was great uncertainty about the extent to which Trump would dismantle environmental protection.

Figure 2: EnvPU - An index of environmental policy uncertainty 1990-2019

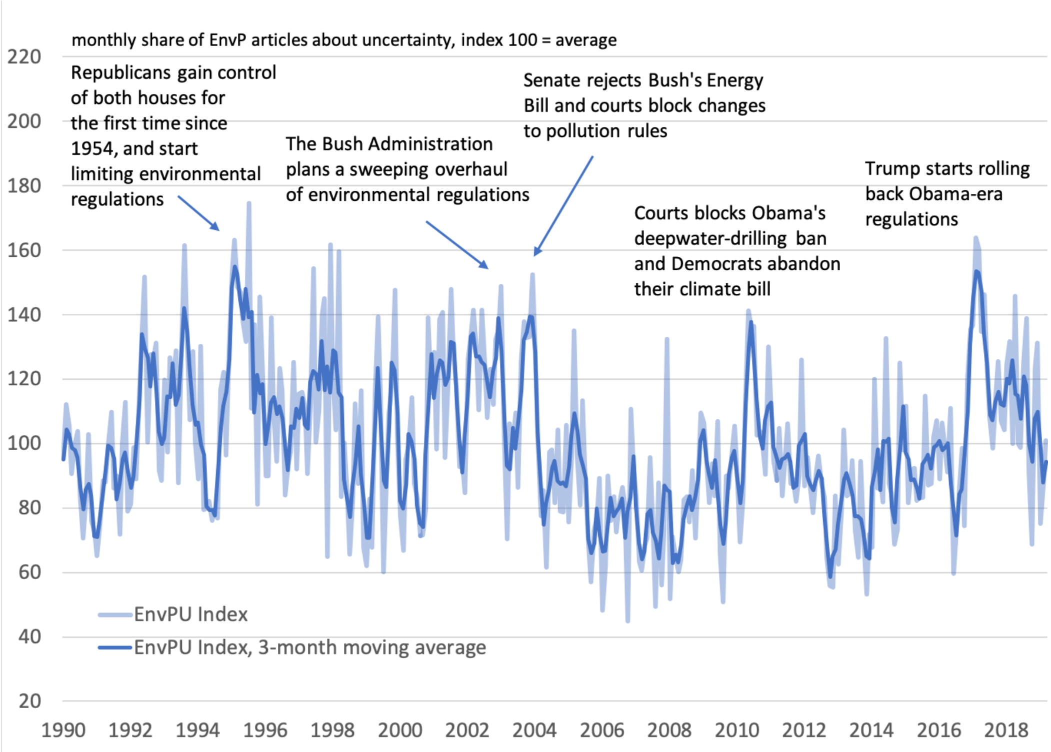
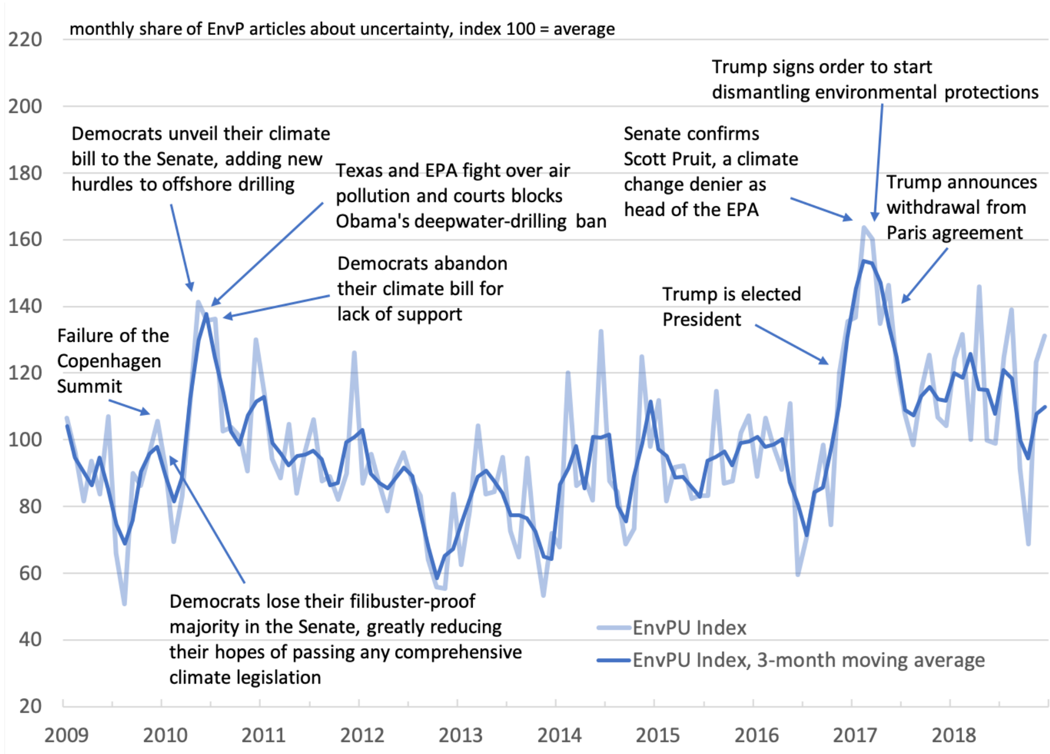


Figure 3: EnvPU - An index of environmental policy uncertainty 2009-2019



3 Evaluating our environmental policy uncertainty index

3.1 Comparison with a keyword-based approach

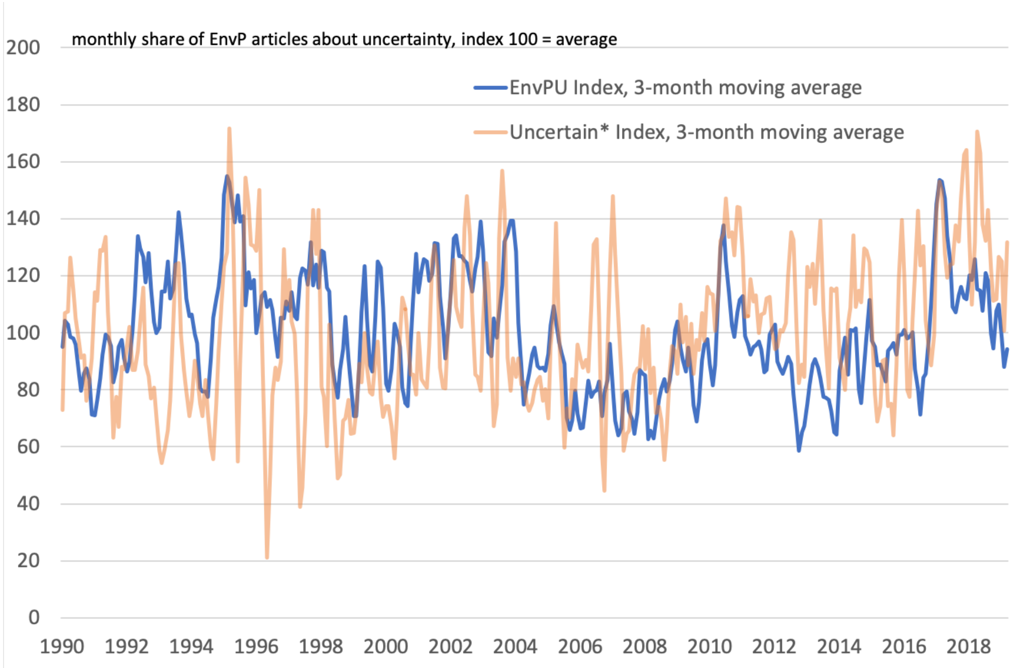
To benchmark our supervised machine learning index, we compare it with a 'naive' dictionary-approach which consists in searching for environmental policy articles including *uncertain**[8] keywords as in Baker et al. (2016). We count the number of articles with *uncertain** keywords per month in the subset of 80,045 environmental policy news, scaled by the monthly volume of environmental policy news. The index is standardized such that 100 represents the average over the 1990-2019 period. We also discuss in Appendix C a more elaborate classification, based on an extensive dictionary of keywords

[8] i.e. *uncertain*, *uncertainly*, *uncertainty* or *uncertainties*

related to uncertainty. This method does not yield better performance metrics than our SVM algorithm.[9]

Figure 4 plots the 3-months moving average of both the EnvPU and *uncertain** indices. The most striking difference is that the *uncertain** approach yields a much more volatile index oscillating around its average value of 100 and displaying fewer trends.

Figure 4: EnvPU versus *uncertain**, 3-month moving average.



Using our training set as a benchmark, we are able to compute the performance of the *uncertain** approach. Using *uncertain** yields a precision of 49 percent and a recall of 8 percent.[10] While the articles identified as EnvPU by this method will indeed be about environmental policy uncertainty nearly half of the time – lower than our precision of 56 percent – the recall is very low. Indeed, 92 percent of newspapers articles about

[9] Tobback et al. (2018) compare – like we do – three similar methods to reproduce the EPU index from Baker et al. (2016): SVM, naive and a longer list of uncertainty modal words. They find that the latter approach is not significantly better than the first two.

[10] yielding a low F1-score of 14 percent

uncertainty will be missed by this approach. This finding is aligned with Tobback et al. (2018) who compare various algorithms to reproduce the EPU index from Baker et al. (2016) and also find that the 'naive' approach suffers from a very low recall.[11]

Table 2 shows that most articles about environmental policy uncertainty do not use the term *uncertain** but instead use a wide lexicon of words related to uncertainty (e.g., President Trump [...] has reversed course on nearly two dozen environmental rules [...]). This comparison sheds new light on the downsides of the keyword approach to capture policy uncertainty. While it is true that simple *uncertain** keywords can reliably identify some of the articles about policy uncertainty it will miss the vast majority of them, which begs the question of whether some topics are systematically omitted. Additionally, being entirely dependent on very few terms increases the volatility of the *uncertain** index, as Figure 4 illustrates. Our machine learning approach, because it is based on a wide array of features, is able to cast a much broader picture. As a result, it identifies many more articles about policy uncertainty than the *uncertain** approach, without weighing on its precision.

3.2 Human Audit

To further validate the EnvPU index, we perform an additional human audit study. Humans may inevitably disagree on what type of article or wording reflects uncertainty and what does not — and we certainly cannot expect our classifier to do better than our annotators. However, some level of inaccuracy may be acceptable so long as both our SVM and human-based approach identify the same trends in policy uncertainty.

By providing us with a human-based index that we can compare to the machine-based indices (i.e., SVM and *uncertain** approach), we aim to verify that our computer-based index does not miss any significant change in the level of policy uncertainty. To

[11]In their case, a SVM algorithm produces a recall of 68 percent, while the naive approach has a recall of 21 percent. Precision metrics are higher (88 percent and 70 percent respectively) with both methods, as they rely on a one-step approach searching for economic policy uncertainty in all news - leading to patterns easier to identify for the algorithm - rather than a two-step approach as we do.

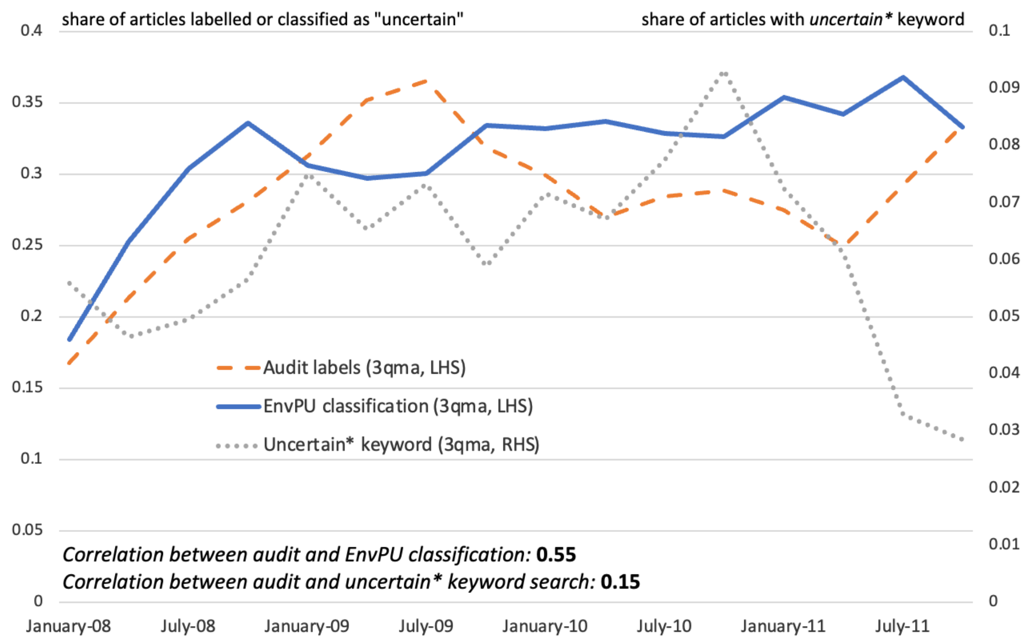
that end, we hired six human auditors to read and label a sub-sample of 925 articles randomly drawn from all the 14,158 published articles over the January 2008-December 2011 period (i.e., around 6.5 percent) in batches of overlapping samples. We choose this period because it is a central period in terms of US environmental policy uncertainty, with Obama taking over as president and pushing a green policy agenda. Unaware of the classifier's label, the annotators manually labeled each article in pairs of two based on our codebook and continuous training we provided. Using the audit labels, we are able to verify the validity of our index and how it compares to the performance of the *uncertain** approach.

Our human audit study reveals that human annotators agree with the SVM label 72 percent of the time. There is still a significant level of disagreement, which seems to be a feature of the task as we also found earlier that annotators tend to disagree about 30% of the time. Our algorithm, as well as any imaginable approach, will thus also make mistakes.

Most importantly, we investigate whether our SVM index picks up the same trends as our human-based index and how this compares to the *uncertain** keyword approach. Figure 5 plots the 3-quarters moving averages of these three indices based on the number of articles identified as relevant on a quarterly basis. While the human- and SVM EnvPU indices capture the same overall trend of increasing uncertainty during 2008 followed by a small decrease, with an uptick in 2011, this trend is less visible when using the *uncertain** keyword approach. Furthermore, the correlation between the human-based and SVM index (0.55) is significantly higher than between the keyword-based and the human-based index (0.15).

Finally, we test whether our machine learning algorithm makes systematic errors that are influenced by outside events such as the business cycle. To test this, we compute the difference between the number of articles labelled as uncertain by the human auditors and the algorithm and look at the correlation between this difference and a measure

Figure 5: The audit versus the EnvPU and uncertain* approach



of economic growth (i.e., the quarter-on-quarter GDP growth from the U.S. Bureau of Economic Analysis). With a correlation coefficient of -0.13 over 2008-2011, we are confident that the errors of our classifier are not correlated with economic activity.[12]

3.3 Political slant in newspapers

A newspaper-based measure of policy uncertainty is likely to be influenced by the political slant of the newspapers in our sample. Conservative and liberal-leaning newspapers might overemphasize uncertainty when it is caused by their political opponents' action. To study whether political slant influences our index, we divide the newspapers in our sample into two groups based on whether they are more left or right leaning.[13]

[12] See Figure D.1 in Appendix D.

[13] To determine whether a newspaper is more conservative or liberal leaning, we use two external sources: Boston University (https://library.bu.edu/c.php?g=617120p=4452935) and AllSides, a multi-partisan organisation that studies media bias (https://www.allsides.com/).

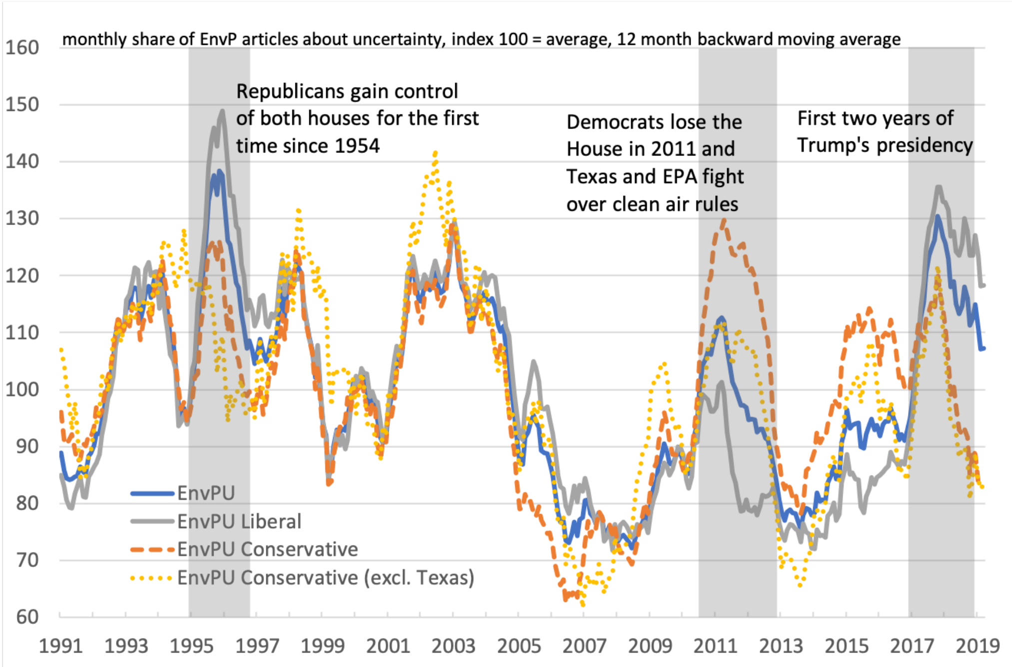
- Liberal-leaning: New York Times, Washington Post, San Francisco Chronicle, Tampa Bay Times, San Diego Union Tribune and San Jose Mercury News.
- Conservative-leaning: Wall Street Journal, Houston Chronicle, Boston Herald and Dallas Morning News.

We plot the EnvPU indices produced by the Liberal-leaning and Conservative-leaning newspapers in our index in Figure 6. An interesting observation is that, while the indices are very similar most of the time, they tend to diverge during periods of high political uncertainty. Left-leaning journalists describes environmental policy in very uncertain terms when Republicans suddenly regain control of the political apparatus while the more right-leaning newspapers are less affected. The most striking example of this occurs in the last years of our index. On the one hand, right-leaning newspapers report the same level of uncertainty when Trump takes office than during the last two years of Obama's presidency. If anything, they describe environmental policy as being less uncertain under Trump. Indeed, at the start of 2019, the level of uncertainty reported by conservative newspapers is at one of its lowest level of the past three decades. On the other hand, left-leaning newspapers report a near doubling of policy uncertainty when Trump is elected, to its second highest level of the past three decades. Moreover, policy uncertainty remains high throughout Trump's early presidency. A similar scenario can be observed in 1995 when Republicans regain full control of Congress.

Liberal-leaning newspapers are not the only ones to emphasize uncertainty when their ideological opponents, Republicans, are in power. Indeed, conservative media outlets reports an increase in policy uncertainty in the last two years of Obama's presidency, when the Paris Agreement is signed, while the more left-leaning newspapers barely register any change in uncertainty.

Overall, the fact that political slant skews the reporting of policy uncertainty is interesting but does not undermine our index. Indeed, our sample of newspapers is well

Figure 6: EnvPU according to liberal and conservative media



balanced between liberal and conservative prints.[14] As a result, our EnvPU index evens out the biases from each side.

Figure 6 also highlights the need to have a geographically diverse set of newspapers. Indeed, in the 2011-2012 period, the Dallas Morning News and Houston Chronicles report a high level of policy uncertainty while other newspapers are less affected. This is due to the fact that the source of this uncertainty, a fight between Texas and the EPA over clean air rules, happens in the home state of these newspapers. Journalists that are directly impacted or exposed to a policy dispute might overemphasize the sense of uncertainty.

[14]The articles from conservative-leaning newspapers actually represent 37 percent of articles in our EnvP sample. While this is not perfectly balanced, it reflects the actual American context where a majority of journalists, and by extension newspapers, identify as liberals (Hassell et al., 2020).

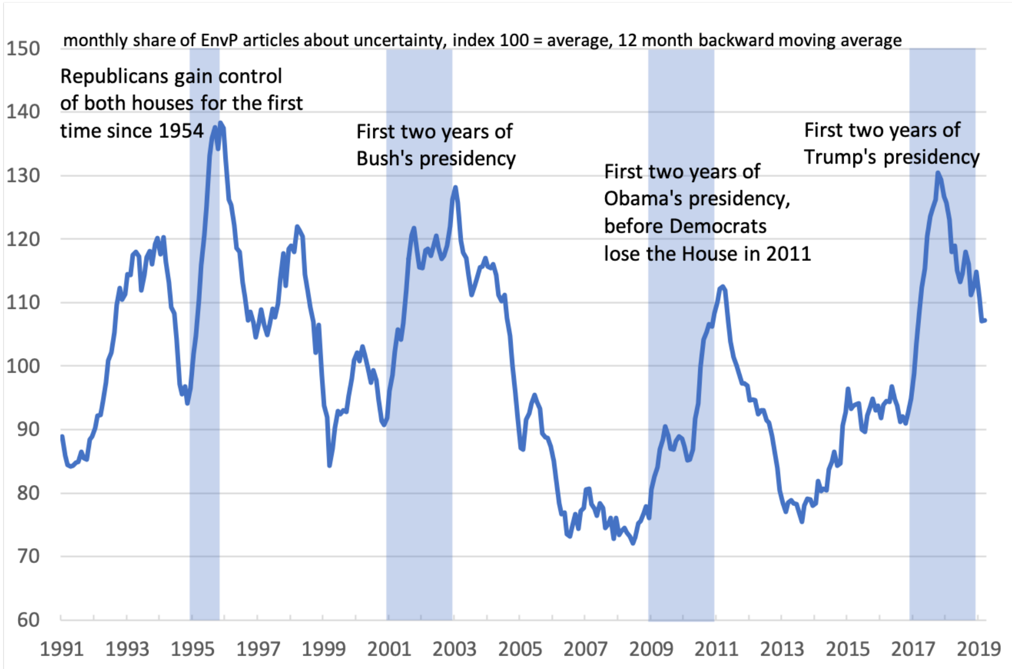
3.4 Environmental policy uncertainty and US elections

In this section, we examine patterns in our EnvPU index around US Congress and presidential elections, as these are of primary importance to make projections about the future environmental and climate policy regime.

We first provide descriptive evidence in Figure 7 that the EnvPU index rises during periods of transitions in US politics. Specifically, the index spikes when Republicans retake control over important policy-making institutions. This is not surprising given that the Republican policy agenda usually includes sweeping roll-backs of environmental regulations, causing a period of rising uncertainty about the future state of environmental policy. Figure 7 shows that the first spike in our index occurs between January and March 1995 when the Republicans take back both houses of Congress for the first time in 40 years and launch an attack on environmental regulations. As noted earlier, levels of EnvPU then skyrocket during the government shutdown at the end of the year 1995. Other notable high levels of policy uncertainty come about in 2003 when the Republicans control once again both houses and in 2017 when President Trump takes office with a clear agenda of rolling back Obama-era regulations.

Another interesting observation is that, after the transition of powers took place, our EnvPU index usually subsides at a low level. A notable illustration of this phenomenon can be found during the Trump era. The EnvPU index stands at around 80, 20 percent below its average level, just before President Trump is elected and takes office at the end of 2016. EnvPU then nearly doubles to 160 in 2017, one of its highest level of the past three decades. However, once the dust has settled, policy uncertainty starts to creep back down, nearly reaching 80 once again in early 2019. This seems to suggest that a low but predictable level of environmental regulations, i.e., when Republicans are firmly in power, translates into a low level of environmental policy uncertainty. This is encouraging because it provides evidence that our index captures uncertainty and not simply the level/stringency of environmental policies.

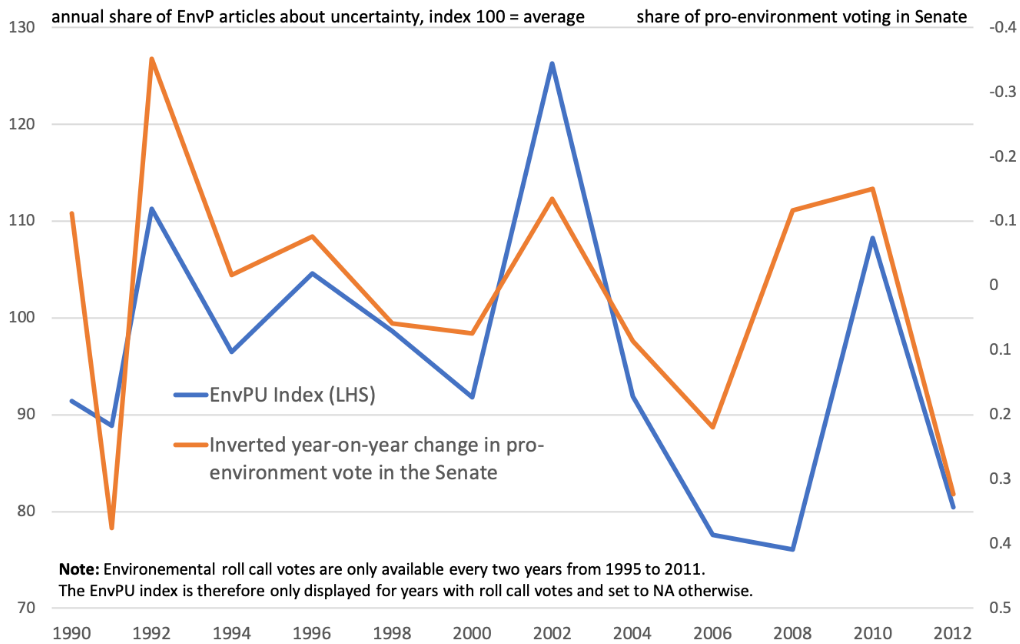
Figure 7: EnvPU during transitional periods in American politics, 12-month moving average



To further explore the links between Congressional elections and EnvPU, we test the relationship between the EnvPU index and changes in the pro-environment composition of the U.S legislature. We use data on environmental roll call votes for the U.S. Congress from 1990 to 2013 (Kim and Urpelainen, 2017). This data allows us to compute the annual share of pro-environment votes for each senator. Taking the average across all senators yields the average share of pro-environment votes in the U.S. Senate. Changes in this average from one year to the next reflect changing party allocations and changing personal opinions. We argue that because our EnvPU index is sensitive to transfers of political power, it reacts to year-on-year changes in pro-environment votes rather than its level. In particular, sudden decreases in pro-environment votes could be more strongly associated with high levels of policy uncertainty at it becomes difficult to project which regulations will be scrapped. A sudden increase in environmental votes could have a more muted effect. Indeed, the prospect of a more supportive environment could dampen the uncertainty, and uncertainty to the upside might be described using a less potent vocabulary.

Figure 8 plots our EnvPU index at an annual frequency and the year-on-year change in the pro-environment votes in the Senate. We see that when the share of pro-environment votes drops in 1992-1993, 2002 or 2009-2010 the level of environmental policy uncertainty spikes. Increases in pro-environment votes also correlate with lower EnvPU. Overall, with a correlation coefficient of 0.6 over the years 1990 - 2013, our EnvPU index is associated with changes in the political composition of the American Congress. In particular, transfers of power to Republicans lead to spikes in environmental policy uncertainty.

Figure 8: EnvPU and environmental roll call votes



As a next step, we study more formally how our EnvPU index fluctuates around major presidential elections. Over our sample period of 1981-2019, we observe nine election cycles, starting with the 1984 presidential election between Reagan and Mondale up to the 2016 presidential election between Trump and Clinton.[15] We define each of the election cycles as the 22 months before and the 25 months after the election, as well as the month of the election itself. In a similar fashion to Baker et al. (2020), we then characterize the evolution of EnvPU in the months around the nine presidential elections by running the following regression:

$$ln(EnvPU_t) = \gamma_m + \gamma_c + \sum_{n=-6}^{6} \beta_n 1(ElectionMonth_{t-n} = 1) + \epsilon_t \quad (1)$$

where t indexes the monthly dates, γ_m is a month fixed effect that deals with the potential seasonality in EnvPU and γ_c is an election cycle fixed effect. Each of the thirteen β_n coefficients captures the level of $ln(EnvPU)$ during the n months surrounding the election relative to the average level of policy uncertainty during its election cycle, everything else equal.

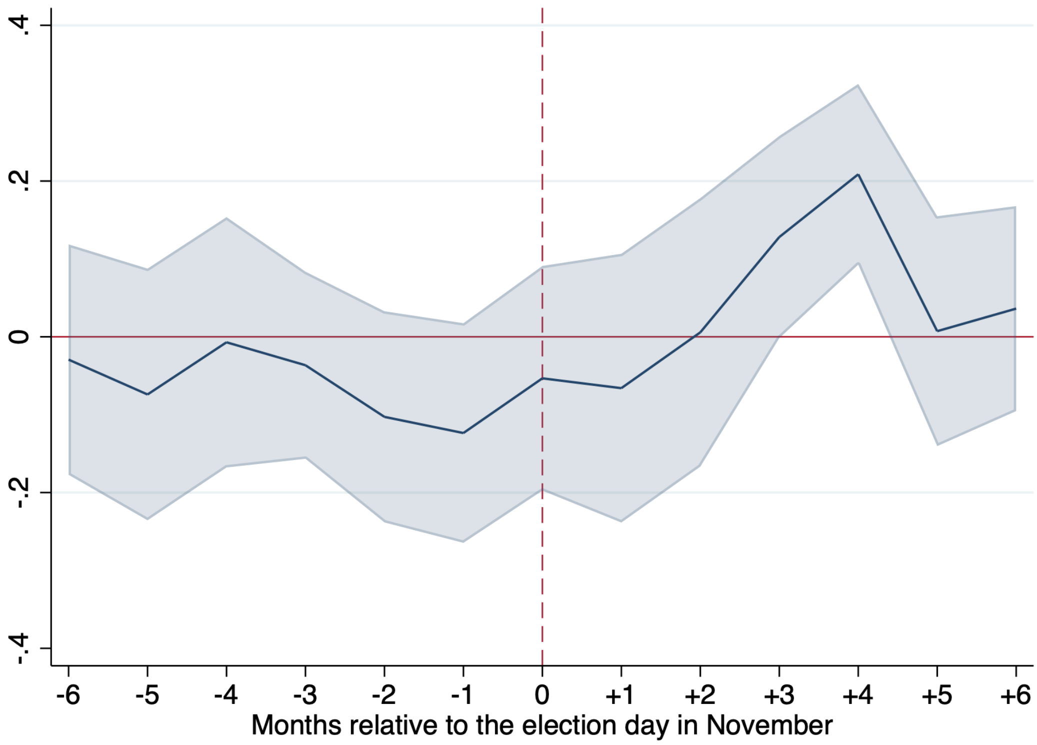
The estimates are in column (1) of Table E.2 in Appendix E. To better visualize the behavior of EnvPU, Figure 9 displays the β_n coefficients for the six months before and after a typical presidential election. This Figure shows that policy uncertainty is slightly lower in the months directly before the election while it is around 21 percent higher four months after the election. The low level of disagreement - and thus low volatility - on environmental policy topics before presidential elections has been documented by McAlexander and Urpelainen (2020). They show that both Republican and Democrat legislators are more likely to cast a pro-environment vote in the 60 days prior to an election. The rationale is that election time is when citizens pay the most attention to legislators, giving the latter incentives to engage in visibly pro-environment behavior.[16] Furthermore, the increase in EnvPU four months after the election is consistent with

[15] Table E.1 in Appendix E details our nine election cycles.

[16] The public prefers stronger environmental policies than the average interest group (i.e., industries that might suffer from regulations).

our previous discussion. It is indeed typically during the first few months of their terms that presidents announce their specific plans (e.g. intentions to rollback or implement environmental regulations). These announcements and the subsequent actions cause policy uncertainty to rise.

Figure 9: EnvPU and presidential elections



This figure presents the coefficients on dummies for six months before and after a presidential election using Equation (1) (i.e., Column (1) in Table E.2). The value of these coefficients reflects the level of EnvPU during each of these months relative to the rest of the sample. The shaded area depicts the 90% confidence interval.

As a last piece of evidence regarding the relationship between presidential elections and EnvPU, we study whether elevated levels of polarization in environmental public opinion affect the behaviour of EnvPU around elections. We determine whether elections in our sample are polarized using data from Kim and Urpelainen (2017) on the state-level differences in pro-environmental attitudes between Republicans and Democrats.[17] Both Figure E.1 and Table E.2 show that during polarized elections, the swings in EnvPU are more pronounced. Four months after a polarized election, environmental policy uncertainty is around 35 percent higher than during the rest of the election cycle. By contrast, none of the coefficients on the month dummies are significant when elections are not polarized.

To summarize, our analysis of the patterns of environmental and climate policy uncertainty around elections serves to validate the plausibility of our index, as we confirm as expected that fluctuations in our EnvPU index track closely with change of powers in US politics. Spikes in our EnvPU index tend to occur quickly after (and not before) presidential elections, as a number of important changes in environmental regulations are likely to be announced once the president is elected.

4 Environmental policy uncertainty and low-carbon investments

Having developed our EnvPU index, we now turn to the central validation exercise of our analysis, which consists in examining how our EnvPU index relates to investments in the low-carbon economy. We consider two proxies for clean investments, namely venture capital funding and stock volatility, both in firm-level regressions and VAR models.

Conceptually, we expect that a rise in environmental and climate policy uncertainty is always bad news for low-carbon investments, as these tend to be heavily reliant on public policies. Independent of the current level of policy stringency, uncertainty about the future regulatory framework may threaten market opportunities and therefore the profitability of clean investments. We thus expect environmental policy uncertainty to

[17] We aggregate this value at the federal level for each race year in the House of Representatives. An election cycle is polarized if our measure of polarization during the year of the presidential election is above its average value over 1980-2012. The presidential election in 2016 is for instance depicted as polarized. Table E.1 indicates which election cycles are classified as polarized.

be negatively associated with venture capital funding for startups engaged in the low-carbon economy. In financial markets, we expect our EnvPU index to be associated with higher volatility for stock returns of firms active in low-carbon activities.

Measuring the effects of environmental and climate policy uncertainty on investments raises endogeneity concerns. As such, separating environmental policy uncertainty from other types of (omitted) uncertainty such as technological or climate change uncertainty, which likely correlate with investments, is challenging. However, we expect that the correlation between these omitted types of uncertainty and our EnvPU is small – motivated by the fact that we expressly excluded non-policy-related uncertainty in our text classification task. We also consider specifications where we control for generic economic policy uncertainty as measured by the EPU index of Baker et al. (2016). Moreover, potential concerns about reverse causality are mitigated by the fact that variations in our index are, at least partly, driven by (exogenous) political elections.

Our identification strategy relies on a differentiation of firms according to their exposure to environmental policy uncertainty. Specifically, we expect to find a stronger empirical association between our EnvPU index and investments in firms active in the low-carbon economy – those more exposed to environmental policy uncertainty – compared to other firms. We also control for many confounding factors via fixed effects and additional variables. In particular, we control for the EnvP index, i.e. for the volume of environmental policy news, so that we identify the effect of policy uncertainty for a given level of media attention on environmental and climate policy.[18] As an additional step, we complement firm-level estimations with VAR models to illustrate the dynamic relationship between our EnvPU index and low-carbon investments at the aggregate level, potentially capturing additional channels (e.g. entry and exit).

[18] We also show in Noailly et al. (2021) that our EnvP index positively correlates with policy stringency.

4.1 Firm-level estimations

4.1.1 VC investments across industries

We first investigate how the EnvPU index is associated with the probability that a startup will receive venture capital (VC) funding. We expect VC funding to be more responsive to changes in our EnvPU index for firms and startups most exposed to environmental policy uncertainty. To test this, we create three industry-categories based on their differentiated exposure to EnvPU: 1) non-cleantech startups - our EnvPU index should have little direct impact on the fate of startups in sectors like ICT or biotech, 2) non-energy cleantech startups (e.g., pollution filters or recycling) - as their business model depends in part on environmental policy support, we can expect our EnvPU index to impact their investment plans, 3) clean energy startups - due to high capital-intensive and irreversible investments, these startups are likely the most exposed and most affected by environmental policy uncertainty.[19]

In line with the literature, we expect an increase in the EnvPU index to be negatively associated with VC financing for cleantech startups (Tian and Ye, 2017), especially for clean energy startups because of their higher exposure to EnvPU.

To test these predictions, we obtain data on VC funding rounds between January 1998 and March 2019 for US startups from the Crunchbase database and aggregate these funding rounds into a firm-quarter panel dataset.[20] We also extract the firm's industry and founding date as well as all the information related to the funding rounds (i.e. date, amount, series) from Crunchbase. Firms are only included in our analysis when they are active. This means that we remove any firm-quarter observations that occur before a startup's founding date. We also remove observations of startups no longer

[19] Clean energy startups require more capital and time to commercialize than other clean industries because of the need to manufacture and deploy new solar and wind technology (Nanda et al., 2015; Gaddy et al., 2017; Popp, 2017).

[20] We focus on series A to J financing, involving firms founded after 1985. This represents around 75,000 different funding rounds.

seeking funding, either because they have gone bankrupt or because they are mature enough not to require VC funding.[21] Excluding firm-quarter observations containing any missing information — including those where the firm is classified as inactive — we obtain 1,056,221 firm-quarter observations on 35,704 unique startup firms. Table F1 in Appendix F provides summary statistics of the variables in our sample.

Cleantech startups belong to Crunchbase's 'Sustainability' industry group and represent 4 percent of overall VC deals, while clean-energy startups in clean energy, battery, renewable energy, wind energy, energy storage and solar industries represent only 2.4 percent of all VC deals.[22] We estimate whether startups that are classified as cleantech or clean energy are significantly more responsive to our EnvPU index than startups in other sectors using ordinary least squares (OLS) as follows:

$$\begin{aligned} VC_{i,t+s} = &\alpha + \beta_1 EnvPU_t + \beta_2 EnvPU_t \cdot Cleantech_i + \beta_3 Controls_{i,t} \\ &+ \beta_4 TimeTrend_t + \gamma_{quarter/year/industry/state/series} + \epsilon_{i,t} \end{aligned} \tag{2}$$

where i indexes the firm, t the quarter and s represents the number of quarters by which our dependent variable, VC, leads our independent variables—with $s = 1$ being our main specification. We use two different measures of VC investments as our dependent variable: a funding dummy, $Funded$, and the logarithm of the total amount of funding a startup receives during a quarter, $Amount$, conditional on $Funded = 1$. β_1 and β_2 are the coefficients of our two main variables of interest. β_1 captures the impact of $EnvPU_t$ on non-cleantech startups. β_2 on the other hand captures the respective impact of our EnvPU variable on cleantech startups. In some specification we differentiate between clean energy startups and other cleantech (excluding clean energy) startups.

[21] Unless a startup has exited or is registered as 'closed' on Crunchbase, it can be difficult to know whether a startup is inactive. We therefore assume that any firm without funding activity for three consecutive years is inactive after this three-year mark. We do so because 75 percent of acquisitions and 85 percent of IPOs happen in the three years after a startup's last funding round.

[22] A detailed overview of all industries can be accessed here on Crunchbase's website.

We control for the following variables that could be confounding our results. First, we control for the media coverage of environmental policy using our EnvP index, allowing EnvP to have a different impact on cleantech and clean energy startups. We control for media coverage in order to differentiate the effect of policy uncertainty from an increase in policy stringency. We standardize both our EnvPU and EnvP indices to facilitate the interpretation of the coefficients. Moreover, we account for the economic outlook, as policy uncertainty might be less important during an economic crisis, by including the year-on-year growth of U.S. GDP borrowed from the US Bureau of Economic Analysis and the Federal Reserve effective funds rate. We also include the log of the West Texas Intermediate crude oil spot price as it correlates with both environmental policy uncertainty and investment decisions.

We also include a set of variables and fixed effects to absorb variation that is unrelated to environmental policy uncertainty but may nonetheless affect our results, including, firm i's age as well as a time trend, and in some specifications an industry time trend. We also use firm, quarter, year and series funding round fixed effects.[23] The quarter fixed effects are used to account for seasonality in the data.[24] The firm fixed effect control for firm-level unobservables such as firm's performance. The other fixed effects also allow us to control for unobserved variables common to all startups in a given year or funding round. Finally, we cluster standard errors at the startup firm level to correct for potential serial correlations in the error term.

Table 3 presents the results of our regressions using Equation 2. Columns (1) and (2) use the probability of getting funded in the next quarter (Q + 1) as the dependent variable. In column (1), we only differentiate between cleantech and non-cleantech startups. We can see that a rise in our EnvPU index is associated with a reduced probability of receiving funding in the next quarter for cleantech startups. The positive coefficient for

[23]The series funding rounds dummies capture whether the investment is a series A, series B all the way up to Series J.

[24]Additional estimations including quarter-year fixed effects provide similar results.

non-cleantech startups could indicate that a rise in EnvPU makes non-cleantech sectors more attractive to VCs. However, this coefficient is relatively small in size, around one forth the size of the cleantech interaction term, which underlines that the effect is concentrated on cleantech startups. In column (2), we separate between clean energy and other cleantech startups (excluding clean energy). We find that policy uncertainty has a stronger adverse impact on clean energy startups than on other cleantech startups.[25] To illustrate the size of the effect on clean energy investments, a one-standard deviation (sd) increase in environmental policy uncertainty from one quarter to the next would decrease the probability of receiving funding by 0.24 percentage points. While this might seem small, the average probability that a clean energy startup will be funded next quarter in our sample is only 6.4 percent. Therefore a one-sd increase in environmental policy uncertainty is actually associated with a 4 percent decrease in a clean energy startup's probability of receiving funding next quarter.

In column (3), we use the natural logarithm of the amount received in dollars, conditional on having received funding, as the dependent variable. Using this alternative dependent variable confirms that cleantech startups are negatively affected. However, this time all cleantech startups experience the same adverse effect. A one-sd increase in EnvPU is associated with an 5 percent decrease in the amount received by cleantech startups.

Comparing the EnvP coefficients in columns (2) and (3) to their corresponding EnvPU coefficients allows us to benchmark the size of the EnvPU effect. For clean energy startups, a one-sd increase to both the EnvP and EnvPU would lead to lower VC investments, both at the intensive (3) and extensive margin (2). The increased uncertainty thus outweighs the positive effect of the increased environmental policy salience. For non-energy cleantech startups, the effect of a one-sd increase in the EnvP and EnvPU are similar in size. Our findings tell us that policy uncertainty is a significant threat to

[25]Moreover, the sum of the EnvPU index and the cleantech excluding energy interaction term is not statistically significant.

environmental policy's aim to foster clean technology investments.

To further test the validity of our EnvPU index, we do the same analysis with alternative policy uncertainty indices. First, we estimate the relationship between investments and the naive environmental policy uncertainty index based on the uncertain* keyword search. Columns (1) and (2) of Table 4 shows that the naive index has a weaker relationship with cleantech than the EnvPU index. Indeed, in column (1) both cleantech and non-cleantech startups' probability of securing funding are negatively associated with the naive index. More importantly, there are no specific cleantech effect when looking at the amount of funding in column (2). In column (3), we show that our result is robust to the inclusion of Baker et al. (2016)'s Economic Policy Uncertainty (EPU) index. As expected given that it is a general policy uncertainty index, the EPU has an effect on all startups, irrespective of their industry, but no specific cleantech effect. More specifically, a one-sd increase in the EPU is associated with a 0.32 percentage point decrease in the probability of getting funded for all startups, with no particular effect on cleantech startups. This result demonstrates the different types of policy uncertainty captured by the EnvPU index and the more general EPU index.

In conclusion, we find evidence across various specifications that environmental policy uncertainty has an adverse impact on the funding opportunities of cleantech startups, both on the intensive and extensive margin. Moreover, this effect is particularly strong for clean energy startups as these are particularly sensitive to policy uncertainty due to their reliance on public policies and long-term investments. Likewise, media coverage of environmental policy has a stronger positive association with investments in clean energy startups. These results imply that our indicators of environmental policy uncertainty based on newspaper articles account for significant variation in clean technology investments, as measured by venture capital funding. Finally, we show that other indices of policy uncertainty do not display this particular relationship with cleantech startups, which gives credence to our EnvPU index.

Table 3: Baseline results: Relationship between environmental policy uncertainty and VC investment in cleantech

	(1) Funded (Q+1)	(2) Funded (Q+1)	(3) Amount (Q+1)
EnvPU index	0.000957**	0.000960**	0.0241**
	(0.000475)	(0.000475)	(0.0116)
EnvPU x Cleantech	-0.00352***		
	(0.00121)		
EnvPU x Cleantech excl. Energy		-0.00159	-0.0734**
		(0.00116)	(0.0294)
EnvPU x Clean Energy		-0.00338***	-0.0703**
		(0.00122)	(0.0316)
EnvP index	-0.00366***	-0.00370***	-0.0345*
	(0.000807)	(0.000807)	(0.0196)
EnvP x Cleantech	0.00516***		
	(0.00102)		
EnvP x Cleantech excl. Energy		0.00272***	0.0565*
		(0.000982)	(0.0314)
EnvP x Clean Energy		0.00501***	0.0729***
		(0.00106)	(0.0255)
Firm FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Industry-time trend	Yes	Yes	Yes
Series FE	Yes	Yes	Yes
Observations	1056221	1056221	57319
Firms	35637	35637	28297
R^2	0.006	0.006	0.119

Table presents results of an OLS regression. The sample period is January 1998 to March 2019. The dependent variable in Columns (1), and (2) is a dummy variable that indicates whether firm i received VC funding next quarter. In Column (3), the dependent variable is the logarithm of the amount received, conditional on having received funding. The news indices are standardized to a mean of zero and unit standard deviation. Other controls include age, oil price, a time trend, GDP and the fed fund rate. Standard errors are clustered at the company level. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Robustness: Relationship between alternative policy uncertainty indices and VC investment in cleantech

	(1) Funded (Q+1)	(2) Amount (Q+1)	(3) Funded (Q+1)
Naive index	-0.00189***	-0.00931	
	(0.000360)	(0.00799)	
Naive x Cleantech excl. Energy	-0.00190*	0.0325	
	(0.00101)	(0.0250)	
Naive x Clean Energy	-0.00215**	0.00928	
	(0.000978)	(0.0227)	
EnvPU index			0.00143***
			(0.000481)
EnvPU x Cleantech excl. Energy			-0.00155
			(0.00116)
EnvPU x Clean Energy			-0.00333***
			(0.00122)
EPU index			-0.00319***
			(0.000686)
EPU x Cleantech excl. Energy			-0.000503
			(0.000868)
EPU x Clean Energy			-0.000346
			(0.000934)
EnvP controls	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Industry-time trend	Yes	Yes	Yes
Series FE	Yes	Yes	Yes
Observations	1056221	57319	1056221
Firms	35637	28297	35637
R^2	0.006	0.119	0.006

Table presents results of an OLS regression. The sample period is January 1998 to March 2019. The dependent variable in Columns (1) and (3) is a dummy variable that indicates whether firm i received VC funding next quarter. In Column (2), the dependent variable is the logarithm of the amount received, conditional on having received funding. The news indices are standardized to a mean of zero and unit standard deviation. Other controls include age, oil price, a time trend, GDP and the fed fund rate. Standard errors are clustered at the company level. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.1.2 Firm-level stock volatility

Next, we consider the relationship between environmental policy uncertainty as measured by our EnvPU index and firm-level stock volatility. We expect firms most active in low-carbon activities to be particularly exposed to environmental and climate policy uncertainty. In line with previous literature which has shown that policy uncertainty harms investments and leads to higher stock market volatility of exposed firms (Baker et al., 2016; Pástor and Veronesi, 2012), we expect investors to find it harder to assess the profitability of firms actively engaged in low-carbon activities when environmental and climate policies supporting low-carbon markets become more uncertain. Pástor and Veronesi (2012) predict that stock volatility rises when policy uncertainty rises, because an unpredictable policy regime increases the volatility of agents' stochastic discount factor. Uncertainty in the stock market usually translates into higher discount rates because agents care less about the future if they cannot count on future profits. Volatility is induced by agents having different beliefs about the future.

We collect the daily stock price indexes for a sample of 438 firms across 9 industries[26] listed on the US stock exchange from January 2008 to March 2019 from Datastream. We compute daily continuously compounded log returns at the firm level as $r_{i,t} = \ln\left(\frac{p_{i,t}}{p_{i,t-1}}\right)$ and then monthly annualized volatility at the firm level as $\sqrt{252} \times \sigma_{r_{i,t}}$.

Since our EnvP and EnvPU indices may be endogenously affected by market activity or anticipated, we follow Brogaard and Detzel (2015) and consider 'innovations' in both indices by extracting the residuals from an AR(7) and AR(3) model of our monthly series of EnvP and EnvPU, respectively. Standard tests confirm that both series are white noise and have no autocorrelation.[27] We standardize these measures to have unit

[26]Industry classification according to ICB with the following industries: Basic Materials, Consumer Discretionary, Consumer Staples, Energy, Health Care, Industrials, Technology, Telecommunications and Utilities. We exclude observations with negative equity or sales values and observations where growth in total assets was larger than 100 percent in absolute value.

[27]Breusch–Godfrey test for higher-order serial correlation, Durbin's alternative test for serial correlation and the Portmanteau (Q) test for white noise.

standard deviation.

We capture firm-level exposure to environmental and climate policy uncertainty by borrowing data on firm-level green revenues shares from FTSE Russell. The data capture the share of revenues from environmental products and services across relevant business sectors.[28] The average firm in our sample has a green revenues share of about 23 percent. 'Energy', 'Utilities' and 'Consumer Staples' industries exhibit the highest shares of green revenues.

To investigate whether environmental policy uncertainty is indeed associated with a rise in stock market volatility of the greenest firms, we estimate the following regression:

$$\begin{aligned}\log(\sigma_r)_i = &\beta_1 \text{Green Revenues share}_{i,t=y} * \epsilon_{t=m}^{EnvPU} + \beta_2 \text{Green Revenues share}_{i,t=y} * \epsilon_{t=m}^{EnvP} + \\ &\gamma_{\text{i}} + \gamma_{\text{t=m}} + \varepsilon_{i,t=m}\end{aligned}$$

where $\log(\sigma_r)$ is the natural logarithm of the annualized realized volatility of the continuously compounded log returns at the firm level, Green Revenues share$_{i,t=y}$ is either the firm-level average annual green revenues share fixed over the 2008-2019 period or the firm-level pre-sample share. We use fixed green revenues shares so that a firm's exposure to environmental and climate policy is constant over time and not endogenously determined. The γ's are firm and month fixed effects, respectively. We provide summary statistics of all variables used in the regression in Table G2 in Appendix G.

Table 5 shows our results. Environmental policy uncertainty has a significant and sizable positive effect on stock volatility of firms engaged in the low-carbon economy, regardless of whether we use average (AVG) or pre-sample green revenues shares. Quan-

[28] Green revenues are captured by 10 sectors, 64 sub-sectors and 133 micro sectors. The 10 sectors are: energy generation, energy management and efficiency, energy equipment, environmental resources, environmental support services, food & agriculture, transport equipment, transport solutions, water infrastructure & technology, and waste & pollution control. For further details, refer to the FTSE Green Revenues Classification System. In some years, the green revenues share is estimated and a range instead of a point estimate is provided. In this case, we always choose the lower bound of the interval as our estimate of firm green revenues shares.

titatively, a one-standard deviation EnvPU innovation leads to a differential increase in volatility for a firm one standard deviation above the mean in terms of average green revenues share of about 0.5 percent. This effect is significant at the 5 percent level and robust to including industry fixed effects. Moreover, in column (2) we compare the top 10 percent firms in terms of their average green revenues share with the bottom 10 percent and find that their stock volatility rises by 1.3 percent in response to an EnvPU shock. In addition, when using the pre-sample green revenues share as an exposure measure as in column (3), a one-standard deviation EnvPU innovation leads to a volatility increase of about 1 percent. This effect is significant at the 1 percent level and also robust to including industry fixed effects as well as an industry-month time trend. Moreover, columns (4)-(5) show that the effect of the EnvPU remains even when controlling for general Economic Policy Uncertainty (EPU). Finally, in columns (6)-(7) we find no effect when using the naive EnvPU index, further underlining the value-added of our EnvPU index.

These results are in line with the growing literature predicting and finding that stock volatility should rise when policy uncertainty is higher (Pástor and Veronesi, 2012, 2013; Baker et al., 2016). Quantitatively, the effect is in line with but larger than what Baker et al. (2016) find for EPU and stock volatility of highly exposed firms (0.11 percent).

Table 5: Estimation results - EnvPU and Stock Volatility

	(1) Log volatility	(2) Log volatility	(3) Log volatility	(4) Log volatility	(5) Log volatility	(6) Log volatility	(7) Log volatility
EnvPU × AVG GR share	0.0050** (0.0024)			0.0050** (0.0024)			
EnvP × AVG GR share	0.0006 (0.0018)			0.0010 (0.0018)		0.0020 (0.0018)	
EnvPU × Top 10% Green		0.0133* (0.0082)					
EnvP × Top 10% Green		0.0043 (0.0062)					
EnvPU × Pre-sample GR share			0.0097*** (0.0030)		0.0100*** (0.0030)		
EnvP × Pre-sample GR share			-0.0017 (0.0016)		-0.0013 (0.0017)		0.0010 (0.0019)
EPU × AVG GR share				-0.0067*** (0.0015)			
EPU × Pre-sample GR share					-0.0081*** (0.0024)		
Naive EnvPU × AVG GR share						0.0013 (0.0022)	
Naive EnvPU × Pre-sample GR share							0.0045 (0.0032)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39955	39955	17869	39955	17869	39955	17869
Firms	438	438	158	438	158	438	158
R^2	0.65	0.65	0.6401	0.65	0.64	0.65	0.64

The table presents results of an OLS regression. Standard errors are clustered at the firm level. The dependent variable corresponds to the natural logarithm of the annualized monthly volatility of daily log returns. Firm controls include size as the natural logarithm of market capitalization, profitability as return on assets and leverage as total debt over total equity. The EnvP innovations and EnvPU innovations correspond to the residuals from an AR(7) and AR(3) process, respectively, and are standardized to a mean of zero and unit standard deviation. The green revenue share (GR share) is standardized in the same way. The recession associated with the Global Financial Crisis is excluded. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Environmental Policy and Aggregate Clean Investments

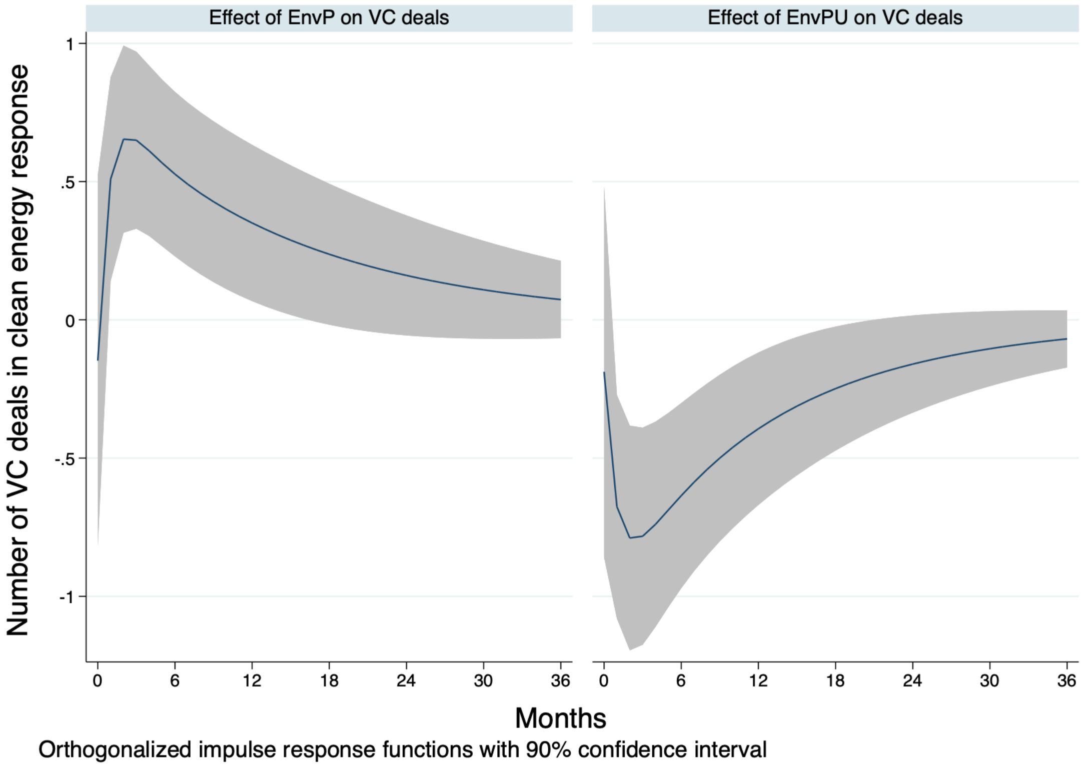
4.2.1 Aggregate cleantech venture capital deals

We now consider the dynamic relationship between policy uncertainty and aggregate venture capital (VC) activity in clean energy technologies. To study this relationship, we use i3 Cleantech Group's data on early-stage financing of cleantech startups, which offers consistent coverage over the past two decades. This database provides information on 11,620 early-stage cleantech deals (seed, series A, series B and growth equity) in the U.S. tracked over time by the Cleantech Group. We extract data on the monthly number of VC deals in the 'energy & power' classification (which includes clean energy generation, efficiency, storage and infrastructure) from January 1998 to March 2019. We focus on VC deals involving clean energy startups because, as discussed in Section 4.1.1, these are the investments that should be most exposed to uncertainty in environmental policy.

Our baseline VAR specification includes the monthly number of clean energy VC deals. Our other variables of interest are our news-based environmental policy uncertainty (EnvPU) and environmental policy (EnvP) indices. We expect VC deals which take several months to close to be more strongly related to the medium-term level of environmental policy uncertainty in the media rather than its specific monthly-value. We, therefore, use the three month backward-looking moving averages of the EnvPU and EnvP indices. We also include the following controls: 1) oil prices as the West Texas Intermediate crude oil spot prices from the St-Louis FED, 2) market risk captured by the Federal Reserve effective funds rate from the Board of Governors of the Federal Reserve System, 3) aggregate economic activity using the Markit's U.S. monthly real GDP index and 4) a linear time trend. We include one lag of all variables, based on lag selection criteria. We conduct standard unit root tests and we use the monthly first difference of the log of oil prices, the log of GDP and the Federal funds rate, because these are not stationary in levels. We keep our EnvPU index in level as it is stationary under all unit root tests. As we can reject the presence of a unit root for the number of VC deals as well as the EnvP index using the Phillips–Perron test, we keep these two variables in levels in our preferred specification, as we are more interested in the level of EnvPU than its month-on-month change. In order to recover orthogonal shocks we use the following Cholesky ordering: EnvPU index, EnvP index, oil price, GDP, the effective Fed funds rate and finally the number of VC deals in clean energy.

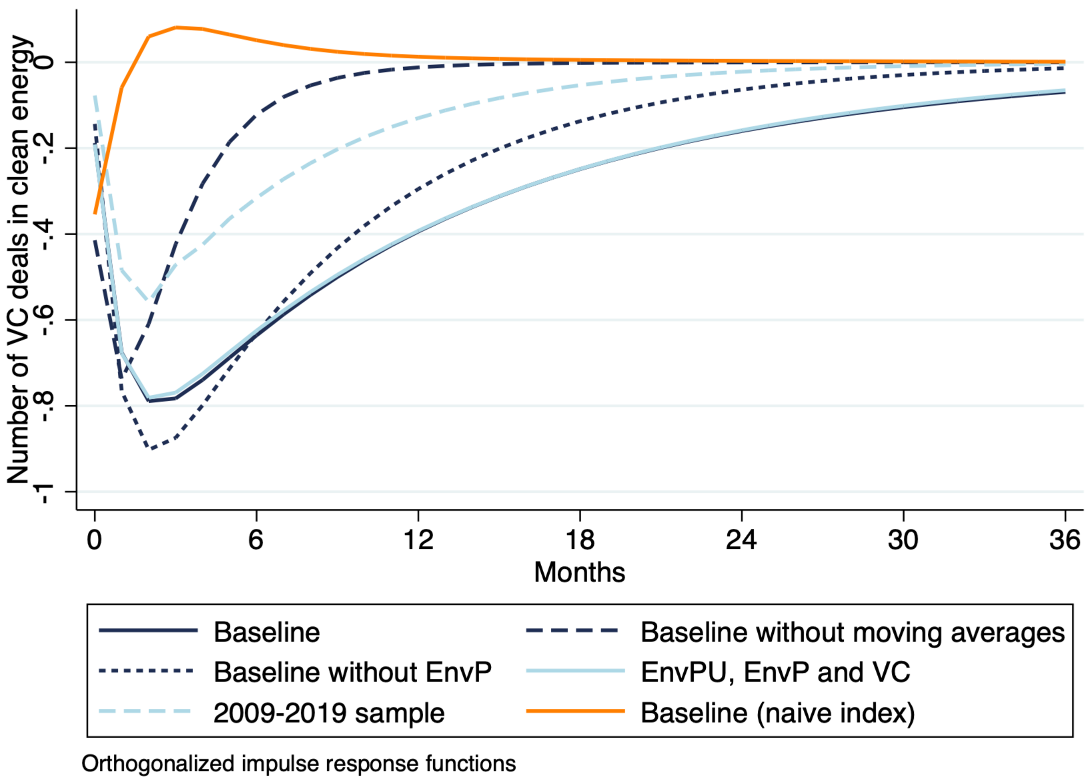
Figure 10 displays the orthogonalized impulse response functions of the number of VC deals in clean energy to both a shock to environmental policy uncertainty and a shock to environmental policy. In the right panel, we see that a one standard deviation increase in environmental policy uncertainty is associated with between 0.5 and 0.8 fewer VC deals in clean energy during the first year after the shock. Conversely, a one standard deviation increase in environmental policy is associated with around 0.5 more VC deals in clean energy (see left panel). While the effect of policy uncertainty is moderate in size, losing half a VC deal still represents a sizable 4.2 percent decrease in the average monthly number of VC deals in clean energy in our sample (i.e. 15.6 between January 1998 and March 2019).

Figure 10: Estimated effect of a one-sd shock in EnvP and EnvPU on the number of clean energy venture capital deals. Impulse response functions to our news-based EnvP and EnvPU indices. The policy indices are smoothed using a three-month backward-looking moving average.



We test the robustness of our results to varying specifications and to using the naive environmental policy uncertainty index instead of our EnvPU index in Figure 11. The negative relationship between EnvPU and VC investments in clean energy holds across specifications. Our baseline response is very similar to the response from a VAR model without our EnvP index, as well as the response without including our controls. Moreover, changing the sample to only cover the post-Great Financial Crisis (July 2009) period or not using moving averages to smooth our indices, does not fundamentally change our results. Interestingly, Figure 11 also shows that the naive index does not display any meaningful dynamic relationship with VC investments, which confirms our previous result that our EnvPU index is more meaningfully related to cleantech investments than one based on a more naive keyword approach.

Figure 11: Clean energy venture capital deals responses to EnvPU Shock (IRFs), under alternative specifications and samples.



4.2.2 Aggregate clean energy stocks

In this section, we investigate the dynamic relationship between our EnvPU index and the volatility of the assets under management (AuM) of the Invesco WilderHill Clean Energy exchange traded fund (PBW-ETF). By the same rationale as before, we expect EnvPU news to raise the volatility of AuM for the PBW-ETF as investors find it harder to predict green firms' future profitability.

Our baseline VAR specification includes: 1) the three-month moving-average of the EnvPU and 2) EnvP index as well as 3) the monthly volatility of daily oil prices, as the US West Texas Intermediate crude oil spot price, 4) the monthly volatility of daily technology stock prices, using the NYSE Arca Technology Index (PSE), and 5) market risk captured by the first difference of the Federal Reserve effective funds rate, and 6) the monthly annualized volatility of the continuously compounded daily PBW-ETW assets under management. We exclude the recession associated with the GFC (December 2007 - June 2009) from the analysis and include one lag of all variables. We include smoothed versions of our policy indices because a perception of elevated uncertainty tends to be one that builds up over successive events (e.g. repeated news reports of opposition to a major renewable energy bill). Apart from major events, actors on the stock market are thus likely to be more influenced by quarterly buildups of policy uncertainty, rather than individual peaks. We provide summary statistics of all variables used in the regression in Table G2 in Appendix G.

Figure 12: Estimated effect of a one-sd EnvPU shock on the volatility of the AuM of the PBW ETF. Impulse response functions of our news-based EnvPU index. The policy indices are smoothed using a three-month backward-looking moving average.

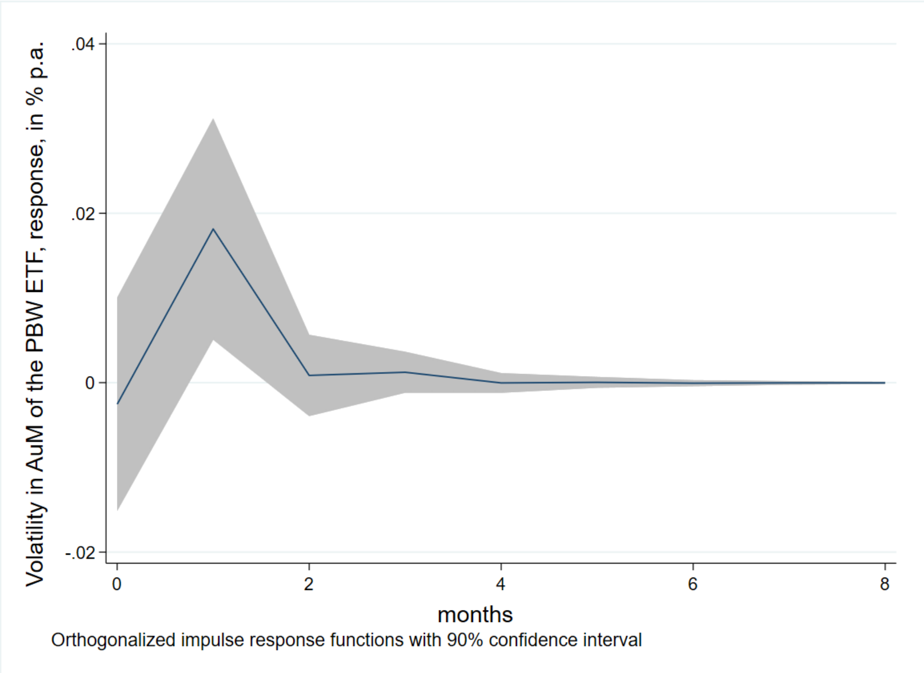
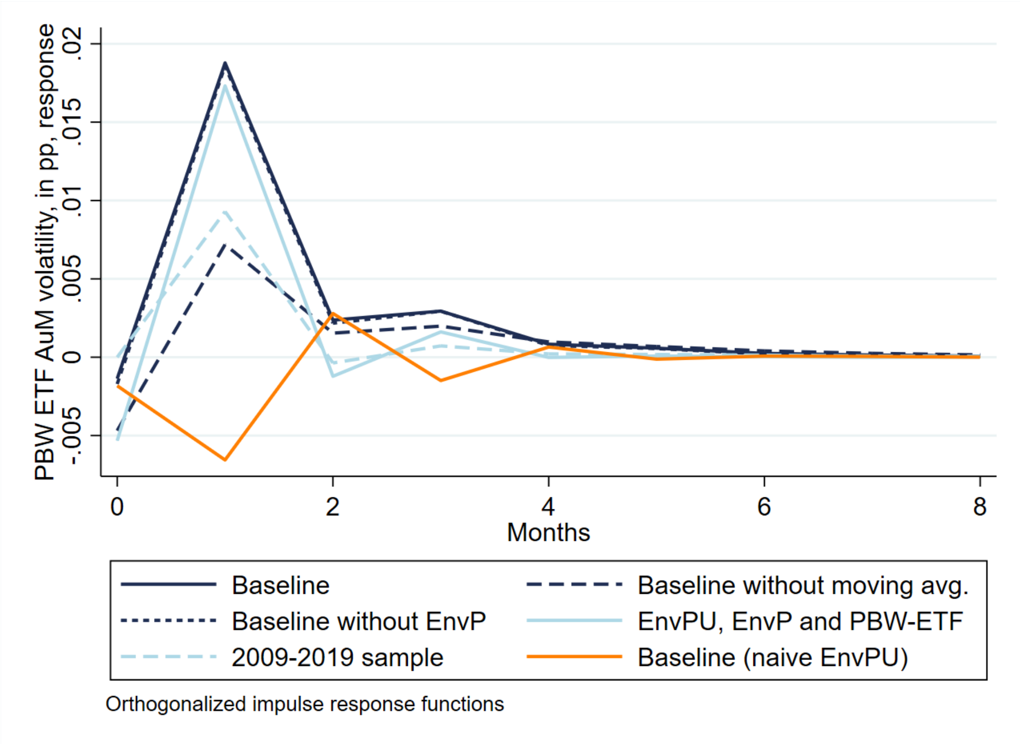


Figure 12 shows our result. A one-standard deviation shock to the growth rate of the EnvPU index leads to an increase in the volatility of the AuM of the PBW ETF of 0.02 percentage points per annum one month after the shock. One reason for this lag may be that ETF investors are largely retail investors who tend to be more passive and less sophisticated in their investment decisions than institutional investors. Therefore, the link between policy uncertainty and investment decision may have more of a medium-term nature.

Figure 13 contains our robustness checks. We test a specification without moving average transformation of our EnvPU and EnvP indices, one without the EnvP index, one with only our main variables of interest, one limiting the sample to the post GFC period (July 2009-March 2019) as well as one using the naive EnvPU index instead of our text mining index. Evidently, the results are weaker when using the non-smoothed policy indices. This is however not surprising given the rationale explained above. Moreover, the effect substantially weakens after the crisis. There is also no effect when using the naive EnvPU index, further underlining the usefulness of our EnvPU index.

Figure 13: Volatility responses of the AuM of the PBW ETF to EnvPU Shock (IRFs), under alternative specifications and samples.



5 Conclusions

A predictable regulatory framework is key to mobilize investments and financial flows towards the low-carbon economy. We apply text-mining techniques on ten leading US newspapers to construct a novel index of environmental policy uncertainty – the EnvPU index – available on a monthly basis over the 1990-2019 period. Our index captures the monthly share of environmental and climate policy uncertainty, scaled by the monthly volume of news on environmental policy. We find that about one-third of environmental policy news report about policy uncertainty, suggesting that the inability to predict how future environmental regulations will unfold is a pervasive attribute of environmental policy discourses. Our EnvPU index correctly captures important spikes in policy uncertainty in the history of US environmental policy, such as the partisan disagreement on environmental spending leading to the 1995-1996 US goverment shutdown, the collapse of the national cap-and-trade policy proposals in 2010 and the environmental policy rollbacks under the Trump's administration as of 2017. We discuss how our novel methodology based on supervised machine learning algorithms outperforms other keyword-based approaches and conduct a series of validity checks, using a human audit and discussing potential bias in newspapers partisan coverage. In addition, we show that fluctuations in our EnvPU index track closely the change of power in US elections, confirming the role of election cycles as a channel of policy uncertainty on environmental and climate regulations.

Next, we examine how our EnvPU index relates to investments in venture capital funding and to the volatility of stock returns of firms engaged in the low-carbon economy. In firm-level estimations, we find that our index is associated with a reduced probability for cleantech startups to receive venture capital funding, especially for clean energy startups characterized by capital-intensive investments that are difficult to reverse. In financial markets, a rise in our EnvPU index is associated with higher stock volatility for firms with greater shares of green revenues. At the macro level, shocks in our index lead to declines in the number of cleantech VC deals and higher volatility of the main benchmark clean energy exchange-traded fund. Altogether, this body of empirical evidence tends to confirm that environmental policy uncertainty threatens the establishment of robust markets for the low-carbon economy, thus slowing the response to climate change.

References

Baker, S., Bloom, N., and Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Baker, S. R., Baksy, A., Bloom, N., Davis, S. J., and Rodden, J. A. (2020). Elections, Political Polarization, and Economic Uncertainty. National Bureau of Economic Research.

Basaglia, P., Carattini, S., Dechezleprêtre, A., and Kruse, T. (2021). Climate policy uncertainty and firms' and investors' behavior. Unpublished manuscript.

Brogaard, J. and Detzel, A. (2015). The Asset-pricing Implications of Government Economic Policy uncertainty. *Management Science*, 61(1):3–18.

Caldara, D. and Iacoviello, M. (2022). Measuring Geopolitical Risk. *American Economic Review*, 112(4):1194–1225.

Caldara, D., Iacoviello, M., Molligo, P., Prestipino, A., and Raffo, A. (2020). The economic effects of trade policy uncertainty. *Journal of Monetary Economics*, 109:38–59.

Dorsey, J. (2019). Waiting for the Courts: Effects of Policy Uncertainty on Pollution and Investment. *Environmental and Resource Economics*, 74(4):1453–1496.

Dugoua, E., Dumas, M., and Noailly, J. (2022). Text-as-data in environmental economics and policy. *Review of Environmental Economics and Policy*, (Summer).

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

European Investment Bank (2021). European firms and climate change 2020/2021: Evidence from the EIB Investment Survey. Report, Luxembourg.

Gaddy, B. E., Sivaram, V., Jones, T. B., and Wayman, L. (2017). Venture Capital and Cleantech: The wrong model for energy innovation. *Energy Policy*, 102:385–395.

Gavriilidis, K. (2021). Measuring Climate Policy Uncertainty. *SSRN Electronic Journal*.

Gentzkow, M., Kelly, B., and Taddy, M. (2017). Text as data. NBER Working Paper No w23276.

Goulder, L. H. (2020). Timing is everything: how economists can better address the urgency of stronger climate policy. *Review of Environmental Economics and Policy*, 14(1):143–156.

Hassell, H. J., Holbein, J. B., and Miles, M. R. (2020). There is no liberal media bias in which news stories political journalists choose to cover. *Science Advances*, 6(14):2–10.

Kim, S. E. and Urpelainen, J. (2017). The Polarization of American Environmental Policy: A Regression Discontinuity Analysis of Senate and House Votes, 1971–2013. *Review of Policy Research*, 34(4):456–484.

Kölbel, J. F., Leippold, M., Rillaerts, J., and Wang, Q. (2020). Ask BERT: How regulatory disclosure of transition and physical climate risks affects the CDS term structure. Swiss Finance Institute Research Paper No. 21-19.

Lemoine, D. (2017). Green expectations: Current effects of anticipated carbon pricing. *Review of Economics and Statistics*, 99(3):499–513.

McAlexander, R. J. and Urpelainen, J. (2020). Elections and Policy Responsiveness: Evidence from Environmental Voting in the U.S. Congress. *Review of Policy Research*, 37(1):39–63.

Nanda, R., Younge, K., and Fleming, L. (2015). Innovation and Entrepreneurship in Renewable Energy. In *The Changing Frontier*, number July, pages 199–232. University of Chicago Press.

Noailly, J., Nowzohour, L., and van den Heuvel, M. (2021). Heard the news? Environmental policy and clean investments. CIES Research Paper 70, Graduate Institute Geneva.

Pástor, and Veronesi, P. (2012). Uncertainty about Government Policy and Stock Prices. *Journal of Finance*, 67(4):1219–1264.

Pástor, and Veronesi, P. (2013). Political Uncertainty and Risk Premia. *Journal of Financial Economics*, 110(3):520–545.

Popp, D. (2017). From science to technology: The value of knowledge from different energy research institutions. *Research Policy*, 46(9):1580–1594.

Sautner, Z., van Lent, L., Vilkov, G., and Zhang, R. (2020). Firm-level Climate Change Exposure. *SSRN Electronic Journal*.

Sen, S. and von Schickfus, M. T. (2020). Climate policy, stranded assets, and investors' expectations. *Journal of Environmental Economics and Management*, 100:102277.

Tian, X. and Ye, K. (2017). How Does Policy Uncertainty Affect Venture Capital? *SSRN Electronic Journal*, (71790591).

Tobback, E., Naudts, H., Daelemans, W., Junqué de Fortuny, E., and Martens, D. (2018). Belgian economic policy uncertainty index: Improvement through text mining. *International Journal of Forecasting*, 34(2):355–365.

A Appendix A: The newspapers in our index

Table A.1: Newspaper distribution of our 80,045 articles

Newspaper	**Available since**	**% of sample**
New York Times	June 1, 1980	22.9%
Washington Post	January 6, 1982	16.2%
Wall Street Journal	June 13, 1979	14.4%
Houston Chronicle	February 2, 1985	13.8%
San Francisco Chronicle	January 4, 1985	8.3%
Tampa Bay Times	June 11, 1986	8.3%
Dallas Morning News	January 18, 1984	6.4%
San Jose Mercury News	January 2, 1994	4.9%
San Diego Union Tribune	December 31, 2010	2.7%
Boston Herald	July 26, 1991	2.1%

B Appendix B: The Support Vector Machine algorithm

Our prediction algorithm relies on support vector machines (SVM). Given a training set $T = ((x_j, y_i), j = 1..n)$ where x_j are the input variables (i.e. the features and their tf-idf score in each article) and y_i is the corresponding output value (i.e. the assigned label for each article), SVM methods will fit a model to this training set, for a given set of parameters. The algorithm builds on the idea that the numerical representation of each text/news are data points in a multivariate space of features. The algorithm aims to find the two parallel hyperplanes separating relevant from irrelevant articles or vectors such that the distance between these hyperplanes, the margin, is maximized. Articles or vectors that lie on one of these hyperplanes, i.e. particularly ambiguous articles that were hard to classify, are called support vectors and the decision boundary is the hyperplane that cuts the margin in half. In other terms, the SVM classifier identifies the most discriminating keywords among the articles which were hardest to classify given the information obtained from our annotations. For example, consider a set of annotated articles where *'EPA'* almost surely indicates the article to be relevant and *'weather'* indicates the opposite. Then, a hypothetical article about the consequences of *EPA* regulation for large *weather* shocks would be hard to classify and might be a support vector. SVM would then give particular importance to words contained in this article that were *not* present in the articles close to the other side of the decision boundary. The rationale here is that less weight can be given to articles further away from the decision boundary as their 'true' (annotator-given) label is simple to infer.

Choosing the optimal hyperparameters

We choose the linear kernel function for its best performance. On a more technical note, we rely on a GridSearch function to set up the hyperparameters adapted to our classification model. This procedure is simply an exhaustive search through a subset of the hyperparameters available for the model (the kernel, the regularization parameter, the penalty parameter, gamma, and the class weight). Using this function we can find the optimal combination of hyperparameter values for our model.

Evaluating the classifier's performance

In order to evaluate the performance of our classifier, we estimate its out-of-sample performance via tenfold cross-validation. After randomly segmenting the training sets into ten sub-samples, the tenfold cross-validation approach consists in estimating the model on nine of the sub-samples and testing its out-of-sample properties on the tenth one. The procedure is then repeated for every possible permutations of the samples. We obtain a quantification of the performance of the algorithm, which is an average over repeated estimations of five ten-fold cross-validations using different random seeds.

C Appendix C: A thesaurus-based index of environmental policy uncertainty.

We construct an extensive list of keywords based words extracted from a variety of sources such as the list of modal words in Tobback et al. (2018), our own set of manually labelled newspaper articles and other sources that led to a list of 270 unigrams and 160 bigrams. The main objective of this exercise is to investigate whether a more sophisticated keyword-based approach to identifying environmental policy uncertainty – arguably a fairer comparison to our more complex machine-learning algorithm – can do better than the 'uncertain*' approach commonly used in the literature. To do so, we collect a list of unigrams and bigrams which reflect our concept of environmental policy uncertainty. To mitigate overfitting, we do not select keywords based on their performance in the training set.[29] The query is made up of keywords shown in Table C.1.

Performance of the Thesaurus approach

To assess the overall performance of our more elaborate query, we generate an uncertainty ratio of how many words of any given article are contained in the query as a fraction of total words in the article. Based on this ratio, we can then define a threshold x above which an article is classified as talking about environmental policy uncertainty (e.g. EnvPU = 1 if uncertainty ratio $\geq$ x%). Next, we compute precision and recall of all classification rules, i.e. for uncertainty thresholds between 0.1% and 3%. This is what Figure C.1 purports to show. Intuitively, precision is highest the higher (more restrictive) the uncertainty threshold is but this comes at the cost of a lower recall. We choose an uncertainty threshold of 0.4% for our baseline prediction rule with a precision

[29] In this context, overfitting can become a problem if we used the same articles as a source of suitable keywords *and* as a test set to assess the performance of the set of keywords. In this case, the performance metrics would be biased upwards because the test set is no longer random but has been used to inform the choice of keywords.

Table C.1: Keywords for predicting our EnvPU index

Unigrams	Bigrams	
ambivalence	against compromise	no clarity
battle	amid skepticism	no clear
challenge	angry talks	not credible
clash	appear slim	not resolved
delay	awaiting action	not settled
disagreement	back away	oppose bill
divide	biggest rift	oppose proposals
divisions	block agency	oppose renewal
expire	change drastically	protracted battle
loopholes	constitutional challenge	pushing reauthorize
obliterate	contentious issue	rival proposals
obscure	court appeal	roll back
obstruct	court order	sidestep epa
override	deep divisions	significant opposition
overrule	deeply split	surprise twist
overthrow	difficult challenge	uncertain outcome
overturn	dramatic steps	uphill battle
pending	due expire	veto bill
polarizing	ever attempted	vowed sue
postpone	extremely complex	is uncertain
repudiate	fate uncertain	higher uncertainty
reverse	fiercely oppose	
rift	fight court	
setback	filed suit	
stall	grave concerns	
sue	hard overcome	
tentative	holding up	
unanticipated	hot debate	
vague	intense battle	
lawsuit	legal challenges	
court	negotiation impasse	

of 0.55, recall of 0.53 and an F1 measure of 0.54. These numbers represent a significant improvement compared to the performance of the ‘uncertain*’ query. However, they are still inferior as compared to our SVM based algorithm.

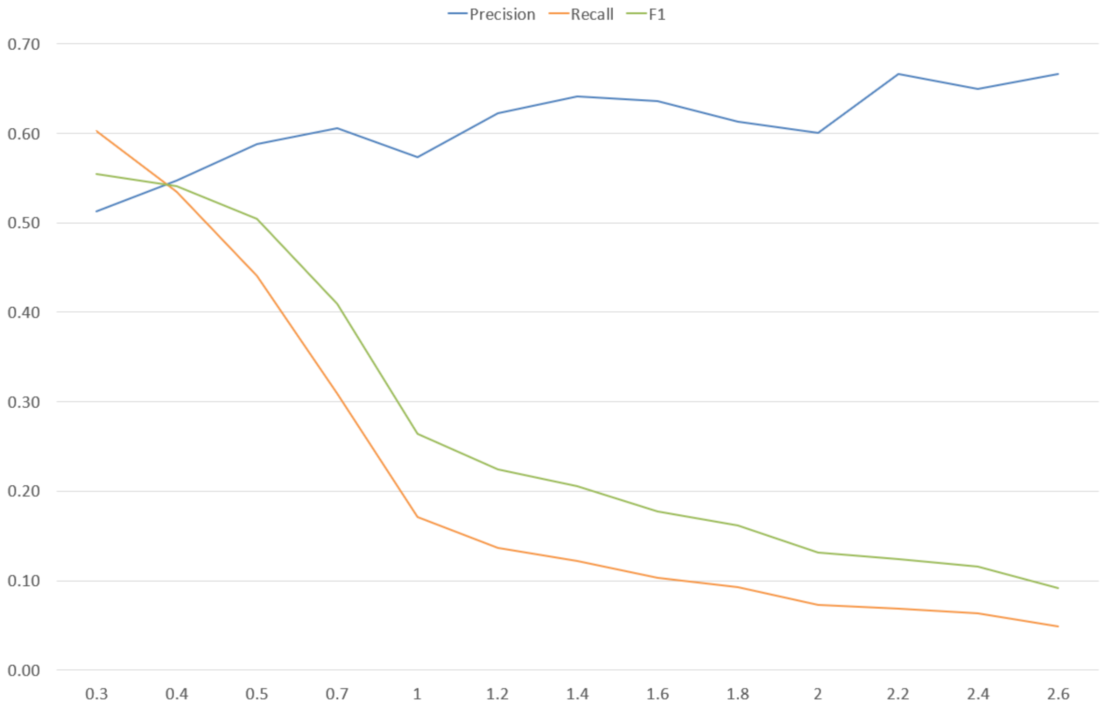
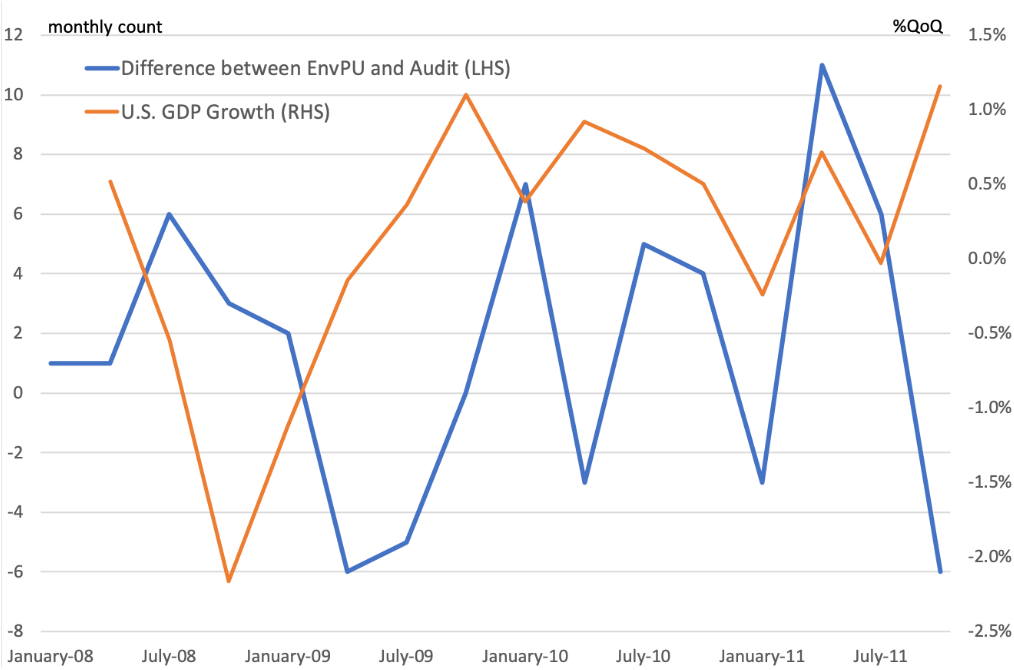


Figure C.1: Query performance for different uncertainty thresholds

D Appendix D: Extra material on the audit

Figure D.1: The difference between the Audit and EnvPU versus GDP growth



E Appendix E: EnvPU and elections

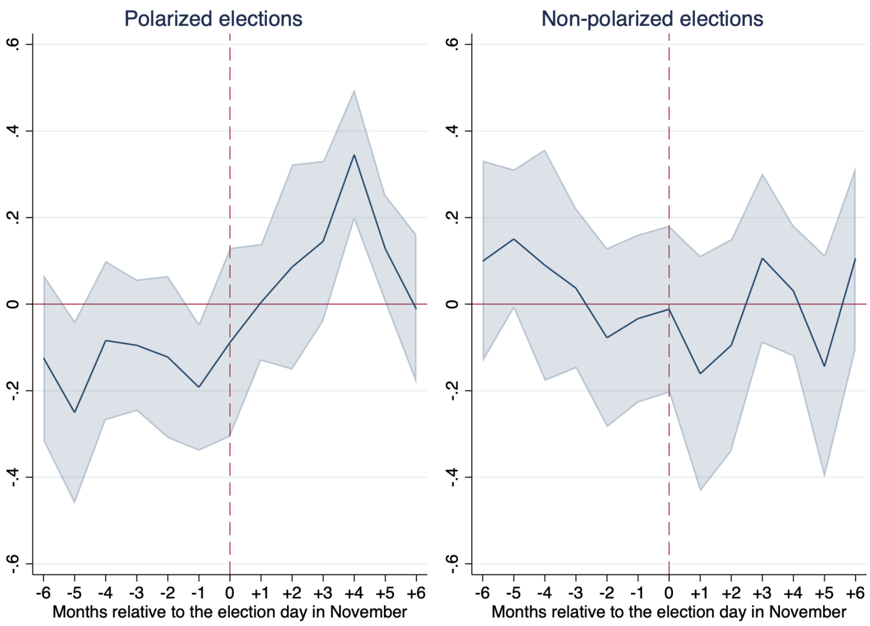
Table E.1: US Presidential Elections, 1984 onwards

Election	President elected	Polarized
1984	Ronald Reagan (R)	No
1988	George H.W. Bush (R)	No
1992	William Clinton (D)	No
1996	William Clinton (D)	No
2000	George W. Bush (R)	Yes
2004	George W. Bush (R)	Yes
2008	Barack Obama (D)	Yes
2012	Barack Obama (D)	Yes*
2016	Donald Trump (R)	Yes**

*: Based on data in two states.

**: Assumed to be polarized.

Figure E.1: EnvPU and presidential elections



This figure presents the coefficients on dummies for six months before and after a presidential election depending on whether the election is polarized (left panel) or not (right panel) (i.e., Columns (2) and (3) in Table E.2). The value of these coefficients reflects the level of EnvPU during each of these months relative to the rest of the sample. The shaded area is the 90% confidence interval.

Table E.2: EnvPU and Elections

		Polarized	Not polarized
	(1)	(2)	(3)
	Log(EnvPU)	Log(EnvPU)	Log(EnvPU)
6 months before election	-0.0292	-0.124	0.0992
	(0.0897)	(0.117)	(0.141)
5 months before election	-0.0740	-0.251*	0.151
	(0.0980)	(0.128)	(0.0972)
4 months before election	-0.00703	-0.0841	0.0899
	(0.0975)	(0.111)	(0.162)
3 months before election	-0.0365	-0.0950	0.0370
	(0.0729)	(0.0919)	(0.112)
2 months before election	-0.103	-0.122	-0.0776
	(0.0822)	(0.113)	(0.125)
1 month before election	-0.124	-0.192**	-0.0333
	(0.0855)	(0.0889)	(0.117)
Election month	-0.0533	-0.0885	-0.0119
	(0.0875)	(0.132)	(0.117)
1 month after election	-0.0659	0.00416	-0.161
	(0.105)	(0.0818)	(0.165)
2 months after election	0.00572	0.0857	-0.0951
	(0.105)	(0.144)	(0.149)
3 months after election	0.128	0.145	0.106
	(0.0787)	(0.112)	(0.119)
4 months after election	0.209***	0.345***	0.0305
	(0.0704)	(0.0916)	(0.0916)
5 months after election	0.00727	0.128*	-0.144
	(0.0895)	(0.0757)	(0.156)
6 months after election	0.0362	-0.0118	0.106
	(0.0798)	(0.104)	(0.128)
Election cycle FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Observations	420	234	186
R^2	.2451552	.284798	.2962674

The table presents results of an OLS regression using Equation (1). The sample period is January 1983 to December 2018. The dependent variable is the logarithm of the EnvPU index. Column (2) restricts the sample to polarized elections and Column (3) to non-polarized elections. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

F Appendix F: Additional statistics and results: Venture Capital

Descriptive statistics - firm-level estimations and VAR

Table F1 reports descriptive summary statistics for the main continuous variables used in our study of VC investments for the period between January 1998 and March 2019. The first panel shows the variables that are common to both the VAR and Panel analysis, taken from our monthly VAR dataset. The second panel displays statistics about the monthly number of VC deals in the United States that we use in our VAR analysis. Finally, the last panel reports the VC-related variables that we use in our panel analysis. The discrepancy between the number of observations reported in the last panel and in Table **??** comes from the fact that, in our analysis, we drop all the observations where the age variable is negative as it most likely indicates an error in the data. In other words we drop around 30,000 observations that were recorded before the startups' official founding date.

Table F1: Summary statistics for venture capital analysis

	Observations	Mean	Std. Dev	Min	Max
Environmental Policy Indices:					
EnvP Index	255	119.75	47.34	40.89	258.55
EnvPU Index	255	96.76	24.83	44.91	163.75
Naive Uncertainty Index	255	101.41	38.04	18.29	235.85
Economic Policy Uncertainty Index	255	111.53	35.47	57.20	245.13
Economic Control Variables:					
YoY GDP Growth	243	2.19	1.75	-4.92	5.64
Oil price (WTI)	255	57.68	28.59	11.35	133.88
Fed Funds Rate	255	2.09	2.11	0.07	6.54
VAR: VC Variables					
Number of clean energy VC deals	255	15.58	11.25	0	48
Panel: VC Variables					
Number of funding rounds, per firm-quarter	1089760	0.06	0.24	0	3
VC amount raised (in mio), if funded	65061	12.16	36.71	0	3500
Age when funded (in years)	63989	5.18	4.40	0	34

G Appendix G: Additional results and statistics on stocks

Table G2: Summary statistics for stock return analysis

	Observations	Mean	Std. Dev	Min	Max
Environmental Policy Innovations					
EnvP Innovations	332	-0.09	23	-75	93
EnvPU Innovations	329	0.2	22	-57	64
Naive EnvPU Innovations	329	0.5	39	-93	142
EPU Innovations	332	-1	30	-115	158
Panel: Stock Policy Exposure Variables					
AVG Green revenue share (%)	39955	22.6	31.4	0	100
Pre-sample GR share (%)	17869	15	32	0	100
Panel: Financial Variables					
Realized stock volatility	39955	0.35	0.29	0.02	15.1
Leverage (debt/equity)	39955	1.6	18.5	0	960.5
Firm size (mln market cap)	39955	15.9	56.2	0.002	1167.2
Profitability (return on assets)	39955	2.5	21.2	-631.5	177.1
VAR: Variables					
Annualized oil price volatility (WTI spot price)	136	0.30	0.13	0.1	0.85
Fed Funds Rate	136	1.13	1.73	0.07	5.26
Annualized tech stock volatility (NYSE Tech 100)	136	0.16	0.07	0.05	0.49
Annualized volatility of the AuM of the PBW ETF	136	0.27	0.1	0.12	0.73

The environmental policy innovations are residuals extracted from an AR(7), AR(3), AR(3) and AR(10), respectively. All financial variables are GDP deflated. The sample excludes the recession period associated with the GFC.