Climate Policy Uncertainty and Directed Technical Change: evidence from European firms

Enrico Bergamini *

January 13, 2025

Abstract

In this paper, I derive novel indexes of Climate Policy Uncertainty for four European countries. Exploiting a new dataset of web-scraped newspaper archives and text-as-data techniques, I explore the role of policy stance underlying aggregate indices of CPU, deriving sub-indexes for uncertainty suggesting increasing or decreasing future stringency. Building on the Directed Technical Change literature, I test empirically the relationship between CPU sub-indexes and environmentally-relevant technologies, in a panel of European firms between 1990 and 2020. I find a significant relationship between the direction of firms' technological efforts, proxied by patents, and that of policy uncertainty. The results suggest that policy uncertainty is a relevant factor in affecting the direction of technical change, bearing important implications in terms of both climate and green industrial policy making.

KEYWORDS: directed technical change; policy uncertainty; green innovation; climate policy; text-as-data

JEL-CODES: O33, Q58, D22

*Department of Economics and Statistics, University of Turin; Collegio Carlo Alberto
email: enrico.bergamini@unito.it

1 Introduction

Addressing the climate crisis, and achieving the Paris Agreement targets of containing average emissions' increase well below 2°C, requires radical decarbonization of the world economy. The scale and speed of change necessary for achieving climate targets entails major coordination efforts, in which the role of governments in steering market forces is fundamental. The scale of these efforts, in fact, has been described as an industrial revolution against a deadline (Schmitz et al., 2013; Lütkenhorst et al., 2014). The development of carbon-neutral technologies in all sectors of the economy, from transport to energy production, plays an essential part in the tension between decarbonization and economic growth (IEA, 2020).

Within the green growth paradigm, the development and production of sustainable technologies open new economic opportunities, while contributing to decouple growth from polluting emissions. The costs required for achieving climate targets and stimulate green innovation must be met timely, in order to avoid further climate damage. Green growth represents an opportunity for the economic success of countries and regions, but is also the source of deep tensions. At its core, climate policy aims at pricing environmental externalities and steering market prices towards making green products and technologies relatively more convenient than polluting ones (Gugler et al., 2024). Transitioning away from a fossil-based model of economic growth is met by resistance of stakeholders of sunset industries. Transitioning away from polluting products and technologies could result in unjust outcomes, by favoring specific economic players, income groups, or territories, and creating new forms of climate-related inequalities (Pegels, 2014; Rodríguez-Pose and Bartalucci, 2023).

In this landscape, the support for climate and green industrial policies critically depends on the success of the policy mix in delivering a just and effective transition (Altenburg and Rodrik, 2017). Climate and green industrial policies have taken a central role in the academic and public debates during the past decade (Wade, 2014; Cherif and Hasanov, 2019). Strong government intervention is essential in steering economies towards a sustainable growth path, and the implications of this necessity are at the core of the tension between the State and the market (Mazzucato, 2011; Rodrik, 2014).

The crucial role of green innovation in ensuring both emissions reduction and competitiveness is everyday more relevant (Fankhauser et al., 2013; Aghion et al., 2023). In this sense, directing technical change away from a high-carbon equilibrium towards a low-carbon one, requires a policy mix able to steer economic incentives for innovation in a cleaner direction (Acemoglu et al., 2012). The need for strong climate policies has been stressed for decades, and progress has been made, but their implementation has been subject to periods of deceleration and doubt. In light of the green-tech race, and more in general the success of the transition, consensus and clarity around climate and industrial policy-making are of immense importance (Altenburg and Rodrik, 2017). Uncertainty in climate policy making has recently been subject to the attention of scholars, as a potential factor slowing down investments and hindering the transition (Basaglia et al., 2021).

As defined by (Baker et al., 2016), Economic Policy Uncertainty (EPU) regards government actions, regulations, and policies that can influence economic and business decisions. Climate policy uncertainty (henceforth, CPU) is a specific subset, focusing on the ambiguity surrounding the design, implementation, or future trajectory of policies aimed at addressing climate change, and achieving the transition. This uncertainty includes unclear government attitudes towards climate regulations, in terms of emissions' targets, carbon pricing mechanisms, or international climate agreements. CPU could in turn affecting the behavior of economic agents, particularly delaying the transition to a low-carbon economy.

In this chapter, I contribute to the empirical literature on directed technical change (DTC), and on the behavior of economic agents facing climate-related uncertainty (Pindyck, 2021), in different respects. First, I build on the empirical literature on policy uncertainty started by Baker et al. (2016), and construct new measures of CPU for France, Germany, Italy and Spain. Exploiting text-as-data techniques and Natural Language Processing (NLP) I derive novel sub-indexes of CPU, leaning towards increasing or decreasing stringency, in order to map the direction of uncertainty. Second, I employ semi-supervised machine learning on this data to make this exercise extensible flexibly to other data sources, and test the relevance of text-as-data techniques in policy-uncertainty and environmental economics applications (Dugoua et al., 2022).

Adding to previous empirical exercises in the literature, I explicitly adopt a DTC framework, and study the effects of CPU on the direction of technological change in firms, studying their low-carbon and polluting patenting activity. I run an empirical analysis on a panel of around 4800 European firms between 1990 and 2020 and argue that the direction of policy uncertainty (suggesting increasing or decreasing probability of future policy stringency) affects the belief revision of firms and in turn the direction of innovation.

The remainder of the paper is organized as follows. Section 2 presents the relevant literature context and develops hypotheses. Section 3 presents the data and the empirical strategy. Section 4 presents and interprets the results. Section 5 concludes and derives policy implications.

2 Literature background

2.1 Policy uncertainty and firms' behavior

In recent years, economists have stressed the importance of endogenous growth in the context of climate change. Starting from Acemoglu et al. (2012), numerous studies have investigated the effects of climate and green industrial policies on the direction of technological change (Dechezleprêtre et al., 2019). A recent body of empirical evidence has confirmed the relevance of DTC frameworks, showing how innovation in climate-relevant technologies is sensitive to economic incentives, and how incentives can be altered by climate policy instruments. (Dechezleprêtre and Hémous, 2022) and Hémous and Olsen (2021) provide recent reviews of the theoretical and empirical developments in this area. Climate policies aimed at affecting the relative price of green and dirty goods, such as carbon taxes or green R&D subsidies affect innovation outcomes, directing technical change away from polluting technologies towards low-carbon ones.

In this context, an emerging stream of literature has started investigating the role of policy uncertainty. The study of firms' behavior in response to uncertainty has a long tradition (Bernanke, 1983; McDonald and Siegel, 1986). More recently, novel forms of climate-related uncertainties are being increasingly recognized as factors affecting the incentives system faced by economic agents (Pindyck, 2021). In addition, the increased availability of text data, in the past decade, has paved the way for a flourishing empirical literature on policy uncertainty, building on the initial idea for EPU proposed in Baker et al. (2016). Differently from other measures of market uncertainty, based often on the volatility of stock prices or on econometric measurement, they developed measures based the text of newspaper articles. Both this work and a large number of follow-up studies have shown the negative effects that EPU shocks exert on the economy during periods of high uncertainty about economic policy actions (for a comprehensive review about measurements and effects, see Cascaldi-Garcia et al. 2023).

EPU indexes are built on a set of keywords able to capture events in which EPU has risen historically, making it possible to operationalize indicators of policy uncertainty across languages and time. Building on similar methodology, an emerging stream of literature has been developing similar indexes for CPU (Gavriilidis, 2021; Basaglia et al., 2021; Noailly et al., 2022). Differently from EPU, CPU is built based on a different set of keywords for climate-related newspaper articles, rather than capturing a broad range of articles dealing with the economy (the precise construction of the index is detailed in Section 3). CPU aims at quantifying how uncertain the climate-policymaking process is, based on a set of nationally relevant newspapers, which might affect the behavior of economic agents.

At the firm level, the responses to uncertainty in adapting expectations are rooted in real-options theory (Dixit and Pindyck, 1994). The effect of CPU on firms' behavior can be understood through two complementary conceptual mechanisms: real-options theory and anticipatory behavior. According to real-options theory, uncertainty about future policies increases the value of delaying investments, particularly when these investments involve high upfront costs or are irreversible (Bernanke, 1983; Dixit and Pindyck, 1994). For green technologies, which are often capital-intensive and highly dependent on regulatory clarity, this mechanism can be particularly relevant. Uncertainty regarding the timing and stringency of measures such as carbon taxes or green subsidies can lead firms to adopt a wait-and-see approach, postponing investments until greater clarity emerges. This delay can be further exacerbated by the path-dependency of green technologies, where early inertia in the development of polluting technologies can create additional barriers to shifting investment priorities.

On the other hand, anticipatory behavior can drive firms to act preemptively in response to policy uncertainty, accelerating investments to gain a strategic advantage in expected future markets. This mechanism may be particularly relevant for green technologies, given their reliance on government intervention to address market failures and their potential for long-term competitiveness in a transitioning economy (Acemoglu et al., 2012) .

While this analysis is built on the notion that firms merely react to CPU, firms may also generate uncertainty by influencing government action by lobbying policy-makers and politicians. While this is highly plausible given the size and relevance of sectors and firms affected by the transition, this avenue of research is beyond the scope of this analysis, and will be further discussed as limitation in my empirical setup, being a source of possible endogeneity. For the scope of the present study, firms are reacting to increasing CPU by a wait-and-see mechanism or anticipatory behaviors, in terms of their technological direction.

Empirically, the net effect of concurrent mechanisms, in the context of climate related risks and uncertainties, is still unclear (Pindyck, 2021). In particular, different studies find rather heterogeneous results, depending on the employed measures for policy uncertainty. In the next subsection, I review the extant empirical literature at the crossroads between policy uncertainty and environmental innovations, and develop the hypothesis.

2.2 Empirical evidence on uncertainty and green innovation

In Table 1, I review of recent studies linking policy uncertainty and firm outcomes, from an environmental and green innovation perspective. I consider two different measures for policy uncertainty. First, I review papers from the literature on EPU, including only studies dealing with environmentally-related outcomes, namely green investments and patenting, or greenhouse gases (GHG) emissions. Second, I include exercises employing CPU as the explanatory variable of interest.

Table 1: Literature Review: Policy uncertainty and innovation

Paper	**Sample**	**Uncertainty**	**Countries**	**Outcome**	**Frequency**	**Direction of effects**
Bai et al. (2023)	firms, 2011-2020	CPU	China	Green Patents	yearly	Positive
Basaglia et al. (2021)	firms, 1990-2019	CPU	US	Stock returns, RD, patenting, employment	quarterly	Negative
Berestycki et al. (2022)	firms, 1990-2018	CPU	12 OECD countries	Investments	yearly	Negative
Bettarelli et al. (2023)	countries and firms, 1976-2020	EPU	81 countries	Green patents	yearly	Negative
Bouri et al. (2022)	firms, 2000-2021	CPU	US	Stock returns (green vs brown)	monthly	Positive
Cui et al. (2023)	firms, 2005-2019	EPU	China	Green Patents	yearly	Negative
Dorsey (2019)	plant, 2002-2011	CAIR (single policy)	US	Investments and emissions	yearly	Negative
Feng and Ma (2024)	firms, 2011-2021	PEU (text-based at firm level)	China	Green Patents	yearly	Negative
Feng and Zheng (2022)	countries, 2000-2022	EPU	22 countries	Renewable Energy patents	yearly	Positive
Gavriilidis (2021)	US, 2000-2021	CPU	US	CO2 emissions	monthly	Negative
Hoang (2022)	firms, 2000-2019	CPU	US	R&D expenditures	quarterly	Negative

Hu et al. (2023)	firms, cross-section	survey-based EnvPU	China	Green investments	yearly	Negative
Huang (2023)	firms, 1987-2019	CPU	US	Green Patents	yearly	Negative
Khalil and Strobel (2023)	macro and firms, 2000-2019	CPU	US	Market value, Investments	quarterly	Positive
Kyaw (2022)	firms, 2002-2020	EPU	US	EnvInnovation score	yearly	Positive
Li et al. (2021)	provinces, 2000-2017	EPU	China	Green Patents	yearly	Negative
Noailly et al. (2022)	macro and firms, 1990-2019	EnvPU	US	Green VC in startups	quarterly	Negative
Peng et al. (2023)	provinces, 2000-2017	EPU	China	Green Patents	yearly	Positive
Ren et al. (2022a)	firms, 2009-2020	CPU	China	Total Factor Productivity	yearly	Negative
Wang et al. (2023)	firms, 2000-2020	CPU	US	CO2 Emissions and Green Patents	yearly	Positive
Wang (2022)	cities and firms, 2003-2019	Local CPU (instrumented)	China	Green RD, Patents, Employment	yearly	Negative
Xu and Yang (2023)	cities, 2005-2016	EPU	China	Green Patents	yearly	Positive
Yu and Chen (2023)	firms, 2007-2020	EPU	China	Green Patents	yearly	Negative

Evidence about the effect of EPU and green innovation is far from conclusive. In a recent working paper, analyzing a large sample of countries and sectors, Bettarelli et al. (2023) suggest that an increase in EPU, measuring the general uncertainty about government's economic policy, depresses green patenting. Cui et al. (2023) and Niu et al. (2023) study the effect of EPU on firm-level green patenting in China, also finding a negative relationship. Li et al. (2021), at the level of Chinese provinces, adds evidence in this direction. Yu and Chen (2023) and Hu et al. (2023) also report a negative association between EPU and green patenting in China, at the firm level. On the contrary, in the United States, Kyaw (2022) and Wang et al. (2023) find a positive effect on measures of green innovation, including investments, patents, and survey-based eco-innovation measures. Xu and Yang (2023) and Peng et al. (2023) finds similar results at the provincial level in China. At the country level, Feng and Zheng (2022) adds to this positive relationship.

EPU is a measure capturing general aspects of economic policy, including monetary policy shocks, terrorist attacks, trade shocks or electoral uncertainty. Environmentally-related technologies might be more sensitive to a general uncertainty shocks compared to other technologies (Bettarelli et al., 2023), because of the different nature of green technologies in terms of risk, complexity, or their need for government support.

Green technologies might be particularly sensitive to policy uncertainty due to their unique characteristics, entailing a higher complexity in their recombinant properties (Barbieri et al., 2020; Fusillo, 2023). Unlike other technologies, green innovations address externalities that markets fail to price adequately, requiring consistent government support through carbon pricing, renewable energy incentives, and green R&D subsidies (Acemoglu et al., 2012). These technologies also involve higher financial and technological risks. Long development cycles and reliance on novel, cross-disciplinary knowledge create significant uncertainty for firms, especially when policy environments are unpredictable. Delays or reversals in key regulations, such as carbon taxes, can therefore devalue investments. Uncertainty could further amplify their exposure to fragmented or inconsistent international policies, disrupting innovation ecosystems and the spatial diffusion of these technologies (Losacker et al., 2023).

Green investments and technologies are therefore arguably more sensitive to uncertainty, and in particular to that specifically bound to climate and environmental policies. An emerging stream of empirical studies, on which this analysis builds, quantifying this type of uncertainty based on the methodology put forward by (Baker et al., 2016). Gavriilidis (2021) measured CPU, based on a sample of nationally-relevant newspapers in the United States, finding a negative relationship with emissions' reduction in a sample of firms. Many studies relate to the effect of CPU rises in the United States, employing the index constructed by Gavriilidis (2021), and observing its relationship with firm level outcomes. Most of these studies focus on the effect of rising CPU in US and Chinese firms. Since then, a number of different exercises have developed alternative CPU indexes.

Noailly et al. (2022) develops a similar index for environmental policy uncertainty

(EnvPU) employing text-as-data techniques, and testing its effect on venture capital funding for US startups. Using quarterly data, they find that an increase in Environmental Policy Uncertainty is associated to lower amounts of capital raised by clean-tech startups. Other recent exercises have built CPU indexes for a larger number of countries. Basaglia et al. (2021) and Berestycki et al. (2022) are the most connected to this paper, measuring and studying the impact of CPU respectively in the United States (Basaglia et al., 2021), and on OECD countries, exploiting a global firm-level datasets (Berestycki et al., 2022). Both studies find a reduction in investments and firm level performances, and Basaglia et al. (2021) also explicitly measures the direction of uncertainty that underlies variation in CPU in English-speaking countries. They use a keywords-based approach to distinguish newspaper articles pointing towards more or less stringent regulation. By interacting emissions intensities as a form of exposure to climate policies, they show how US-firms are more sensitive to variation in CPU that is pointing towards more rigid regulation. While they test results for R&D expenditures, share price volatility, and other outcomes in the US, they do not look explicitly into the direction of effects in terms of green-vs-dirty patenting.

Again for the context of the United States, Wang (2022) measures CPU differently, exploiting the volatility in votes regarding climate legislation, finding that firms adopt an anticipatory behavior with respect to innovation and adoption of climate technologies. Adding to the US evidence, Hoang (2022) distinguishes between low and high-emitting firms, finding a negative effect for the latter, suggesting that heavy-emitting firms might be adopting a wait-and-see investment strategy. Two studies explicitly look at the relative performance of green and dirty outcomes in response to climate policies. At the macro level, Khalil and Strobel (2023) employ both a general equilibrium models, and granular firm level data for the US. They find evidence of capital reallocation towards cleaner assets with respect to more polluting ones, while lowering investments in carbon intensive industries and increasing it in "greener" firms. These findings are in line with Bouri et al. (2022), finding a positive role for US-CPU on the relative performance of green energy stocks vis-à-vis brown counterparts. In a study precedent to the empirical literature based on newspaper data, Dorsey (2019) exploits a quasi-experimental framework relating to a single climate policy measure. He finds that firms exposed to a higher level of CPU reduced investments and experience a lower reduction in emissions.

Outside of the US, a number of empirical studies have investigated the CPU and firms' performance in China. Ren et al. (2022b) find a negative effect of US-CPU on total factor productivity in a sample of Chinese firms, channeled through a reduction in R&D expenditures and cash flows, with results varying according to the institutional ownership of firms. Bai et al. (2023) test the US-based index developed in Gavriilidis (2021) on a sample of listed Chinese firms, finding a positive relationship with green patenting. Similarly, Ren et al. (2022a) find a strong non-linear correlation between CPU and investments, negative in polluting industries and positive for green-related investments. Differently from other studies based on newspaper data, Hu et al. (2023) uses survey-based measures of policy uncertainty (policy content and policy enforcement), at the local level, and find a negative relationship with green patenting in

Chinese firms. Other papers focus instead on Chinese CPU. Recently, Ma et al. (2023) employed deep learning techniques to build indexes with geographical variation of CPU in China. Feng and Ma (2024) also use text analysis techniques on company reports in order to build a measure of environmental uncertainty perceived by the firms, finding that it might hinder green innovation. At the city-level, Wang et al. (2023) finds that a reduction in Chinese green R&D, patents and employment, due to uncertainty specifically constructed around the allocation system of subsidies allocated by the central government.

In summary, while in the case of the EPU index, capturing a more general aspect of policy uncertainty related to economic policies, uncertainty can be expected to be detrimental to any innovation process (Basaglia et al., 2021), depressing general investment activity, the empirical evidence seems to show more mixed results. However, in the case of CPU, the same effect cannot be expected a-priori, as the underlying signals in the climate-policymaking process could be differently affecting environmentally-opposing technologies.

2.3 CPU's direction and technological change

A number of gaps emerge from reviewing the literature on EPU and CPU's effects. First, in terms of technological dynamics, evidence beyond general investments is still scant. Studies focusing on green patenting do not explicitly adopt a directed technical change perspective, controlling for the strong path-dependencies characterizing environmentally-sensitive technologies (Acemoglu et al., 2012; Aghion et al., 2016).

If CPU is differently affecting green and brown investments (or sectors) it is plausible to think that climate-sensitive technologies (green or dirty) would also be affected in different ways. Khalil and Strobel (2023), from a macro perspective in a general equilibrium framework, find evidence for a mechanism of capital reallocation, with investments shifting from brown towards cleaner sectors.

Bouri et al. (2022) add evidence on the positive effect of CPU on the relative performance of green vis-à-vis brown financial stocks. I add to this evidence focusing on technological dynamics. CPU could affect firms' behaviors in terms of future costs and values of the technologies. In an environmentally-positive direction, CPU could rise, for example due to discussion about the implementation of a carbon tax. This would directly affect the (expected) cost of capital for polluting technologies (Khalil and Strobel, 2023), and indirectly the future value of clean-tech alternatives, causing a shift in investments efforts from dirty technologies to green technologies.

From this perspective, policy direction within uncertainty indexes becomes central in the expectation-revision of firms, driving investments towards two alternative technologies. As discussed in depth in the next section, CPU indices are built on a set of environmental, policy, and uncertainty keywords. They capture both directions of the climate policy-making process, aggregating both milestones and setbacks. As mentioned, this aspects could be crucial in terms of firms' expectations and technological trajectories. With the exception of Basaglia et al. (2021) and Berestycki et al.

(2022), most of these studies do not unpack CPU indices, and consider aggregate CPU. Building on their work, I add nuance in terms of the direction of policy-uncertainty. If policy uncertainty points towards a more stringent environmental regulation, firms might be inclined to accelerate innovation in terms of environmental technologies, and divest from fossil-based technologies. A symmetric behavior could be expected in terms of polluting technologies. If CPU increases are driven by setbacks in the climate policy-making process, firms could continue, or accelerate, investments into polluting technologies. Therefore, rather than its aggregate level, a driving factor for innovation could be the underlying variation in "good" or "bad" news for the environment, with uncertainty indicating a higher probability of future regulation that could affect both the costs and returns from alternative technologies.

An increase in positive policy uncertainty would enter the production function of profit-maximizing agents as a (potential) extra cost for the production of environmentally-damaging technologies. A decrease in the probability of a carbon tax, for example, could represent a (relatively) higher expected cost for the development of green technologies vis-à-vis polluting ones.

Hence, uncertainty caused by setbacks in climate policy-making, could cause firms to continue investing in fossil-based technologies. Furthermore, given the strong path-dependency (Aghion et al., 2016), an increase (or a non-decrease) in future value of polluting technologies could be detrimental for development of low-carbon alternative, and incentivize agents to continue developing polluting technologies. In an economic equilibrium which is already favoring polluting technologies, uncertainty could therefore be a significant factor in steering change towards a cleaner path (Acemoglu et al., 2012). By the same logic, an increase in the probability of green subsidies, could decrease the expected costs of firms in developing green technologies with respect to fossil-based ones, therefore incentivizing the former and discouraging the latter. In line with the evidence on green industrial policy (Pegels, 2014), clarity and commitment of legislators around policies is crucial in this sense, as CPU could be an underlying factor altering expectations and investments into alternative technologies. In this chapter, I hypothesize, that different signals underlying CPU matter for the direction of technical change:

- Climate Policy Uncertainty affects the direction of technological change in firms, depending on the underlying changes in the probability of a more stringent environmental regulation.

To the best of my knowledge, no study has yet tested the relationship between (sub) indices of CPU and DTC by considering both low and high-carbon technologies. Furthermore, many of the cited studies using green patents as an outcome variable do not explicitly account for the path-dependency in climate-sensitive technologies, adopting a framework of directed technical change.

The motivation for this study, therefore, stems from the necessity to understand how support for green technologies demands stable, harmonized, and long-term policy frameworks. Without this consistency, the high risks and systemic requirements of green technologies will continue to limit innovation and delay the transition to a sustainable economy. This is particularly relevant in the context of the unpriced externalities and path-dependencies emerging when considering fossil technologies alongside green ones.

I add another two contributions with respect to the extant literature. First, I bring evidence for the European context, while most of the studies reviewed bring forward evidence regarding the US and China. Second, I contribute to the emerging literature applying text-as-data techniques in environmental economics and policy (Dugoua et al., 2022), adopting a novel approach to measure the policy stance of news articles.

3 Empirical Framework

3.1 Data

3.1.1 Newspapers' archives and CPU

In order to build European indices of CPU, I collect millions of full text articles by means of web scraping. Web scraping automates the collection of the content of web pages. I built scrapers for several newspapers archives across my sample countries (Germany, France, Italy and Spain). I focus on multiple archives for each country, as common in the policy uncertainty literature to smooth effects due to the structure of a single outlet.

I collect nationally-relevant archives, although the selection of sources by each country was limited by the availability of digitized news archives and the feasibility of the scraping process. I target outlets with different political leaning, in order to balance reporting biases, which is important for the reliability of this measures, although numerous normalization steps are performed in line with the literature. The selection of newspapers for this paper largely resembles that of similar exercises in the literature (Basaglia et al., 2021).

Table 2: Database of newspaper archives

Country	**Archive**	**Articles**
Germany	Der Spiegel	307,103
	Die Zeit	220,497
France	Figaro	1,773,778
	Le Monde	1,491,681
Italy	Il Foglio	53,541
	La Stampa	5,813,893
	La Repubblica	4,528,482
	Il Sole 24 Ore	149,839
Spain	El Mundo	421,725
	El Pais	2,490,156
Total		17,250,695

With the exception of Germany, for which I focus on the weekly outlets Der Spiegel and Die Zeit, all remaining news sources detailed in Table 2 have a daily frequency. For France, I collect data for Le Monde and Figaro. In Italy for La Stampa, La Repubblica, Il Foglio and il Sole 24 Ore. For Spain, I collect data on El Pais and El Mundo newspaper archives. The resulting dataset allows me to exploit around 17 million full-text newspaper articles, spanning the period 1990-2020.

The data I collect via web scraping include the date on which the article was published, its title, and the body of text. No reliable information about the relevance of the article within that day's newspaper (or within the website) was available. I clean the collected data from duplicates (based on the webpage's URL, its unique identifier). Furthermore I remove near-duplicate texts belonging to different URLs, often resulting from the process of digitization of newspaper scans for articles belonging to the physical editions.

I do not distinguish, in this database, between digitized articles which originally appeared in the paper version, and digitally native articles that gained importance since the early 2000s. While this distinction could help understand the structure of newspapers archives, for most news sources it is not possible to identify the origin of the news article. Thus, I consider the archives available online as a single entity, blending digital and physical news. A more detailed exploration of newspaper data, or a structured digitization of raw scans[1] could have relevant implications for policy uncertainty indicators, but is beyond the scope of this paper. The span and richness of the dataset collected, allows me to build CPU indicators, for four European countries, and a longer time span than the one considered in previous exercises.

Following the methodology proposed by Baker et al. (2016) and subsequent work,

[1]For a recent example see Dell et al. (2024)

I match newspaper articles employing three sets of keywords, and consider an article expressing CPU if it matches all three sets. The first set contains climate-related words (e.g. climate, environment, CO2), for each language. The second set of keywords, instead, matches articles referring to policy issues (e.g. government, policy), including policy-specific terms where relevant (for example ETS - Emissions Trading System). I build on, and expand, the sets of climate and policy keywords adopted in previous exercises (Gavriilidis, 2021; Basaglia et al., 2021; Berestycki et al., 2022).

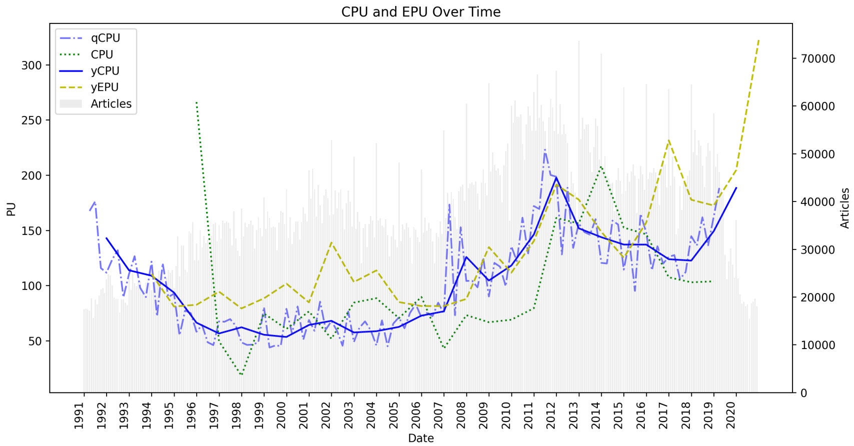
The most important difference, compared to the extant literature, is in the set of keywords expressing uncertainty. The majority of exercises in climate policy uncertainty only match articles based on the keywords "uncertain" or "uncertainty". This selection of keywords has been subject to criticism. Tobback et al. (2018), in the case of EPU, employs a wider set of keywords more generally expressing uncertainty (e.g. doubt, maybe, perhaps). They borrow from the concept of modality in linguistics: modality relates to different ways of expressing degrees of doubts and certainties. Tobback et al. (2018) show that this approach is preferable to simple matches of the words "uncertain" and "uncertainty".

Adapting from this work, I compile a similar list of keywords representing uncertainty. Employing a larger set of keywords helps minimizing false negatives from the sample of articles, by matching a larger number of articles compared to the more restrictive approach. Given this much larger set includes very common keywords, I minimize false positives by only considering articles with a number of modality-expressing words in the top 15th percentile, replicating the approach proposed by Tobback et al. (2018). I match and calculate percentiles separately for each newspaper archive.

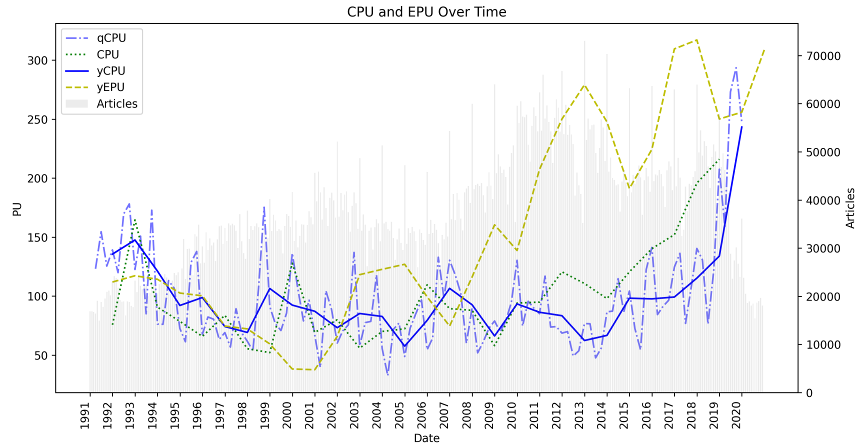
After identifying articles matching the three sets of keywords, I derive monthly time series, for each newspaper, dividing the monthly count of CPU articles by the total amount of articles. In turn, following standard practice in this literature, I normalize the time series by standard deviation, and multiply it for their mean. This normalization helps to remove newspaper-specific factors, due for example to the structure of the newspaper archive.[2] I take averages between newspapers (in the overlapping periods) and multiply the series by 100, making my measures comparable to other policy uncertainty time series.[3]

[2] I calculate standard deviations and means based on periods of consistent number of articles in the archives. I follow Baker et al. (2016), for each newspaper in common in the sample, in defining breaking periods for calculating standard deviations. I also calculate different standard deviations in periods where the total number of articles in the archive shows structural breaks. This might indicate a change in the format of the outlet or in the total amount of digitized news, and could add noise to the measurement.

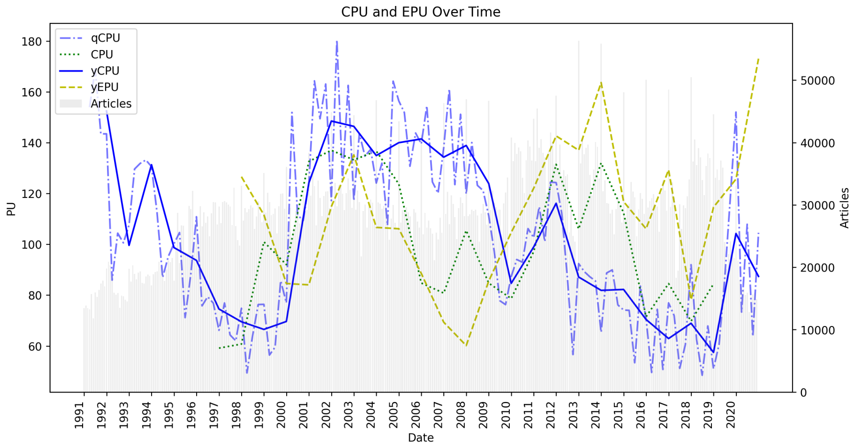
[3] Several series from different exercises are updated and available at: https://policyuncertainty.com



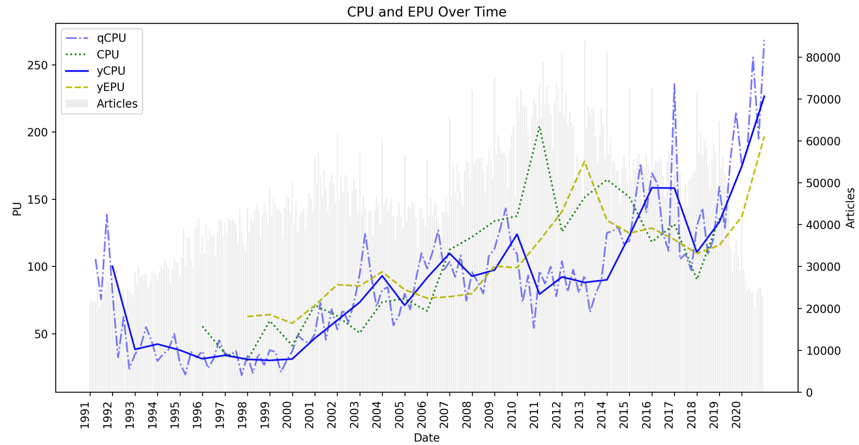
(a) Germany



(b) France



(c) Italy



(d) Spain

Figure 1: Policy uncertainty indexes for sample countries

In Figure 1, I plot these results for the four countries. In blue, I represent the (yearly and quarterly) averages for the CPU index built with the procedure illustrated above. I compare it with a similar index developed by the OECD (Berestycki et al., 2022), as well and the EPU series available from Baker et al. (2016). While it correlates quite highly, in yearly aggregation, with the OECD's CPU index, there are notable differences, most likely due to the different methodologies in keywords matching, and a different selection of newspaper archives. Importantly, the CPU index differs from the Economic Policy Uncertainty.

Interestingly, all indices seem to be spiking around 1992-1993 (years of the discussions around the Kyoto Protocol). Also, spikes in the index correspond to the passing of climate legislation in 2007-2008, when during France's EU presidency, the "Climate and Energy Package" was discussed and adopted, fixing climate targets for 2020. In Germany, the index spikes around 2011, during the discussions on Energiewende, the comprehensive climate-policy agenda for the energy transition, featuring a 60% Greenhouse Gases (GHG) reduction before mid-century. For France, in the early 2000s, some spikes relate to the Climate Act (2001), as well as to the 2003 heatwave (peaking also in Italy and Spain).

In Italy, the index first spikes in the middle of the 1990s, at the beginning of the debate on energy market liberalization, began with the 1996 EU directive, implemented in Italy with the 1999 Bersani Law. Another strong rise happens in Italy, in the early 2000s, during the implementation of the reforms and the privatization of energy markets. Another peak in 2011 refers to peaks in solar panel subsidies.

In more recent years, following 2017, CPU seems to be rising across all four countries. The rising trend starting from 2016, could be due to country-specific factors (the Gilet Jeunes protests following a fuel-tax rise in France) or global discussions. As mentioned, in aggregate, all these indices are summing up different types of events. Both President Trump's decision to exit the Paris Agreement (2017), and the Government's responses to the Climate Strikes began by Gretha Thunberg (2018), or the passing of environmental protection laws could be contributing to the rise of the index. [4]

As noted in Basaglia et al. (2021), while one could expect that a rise in general EPU to be slowing down firms' investments and in general economic activity, and even in green patenting (see Bettarelli et al. 2023 among others), it is not necessarily the same with CPU. As mentioned, the direction of CPU seems particularly relevant for the direction of innovation.

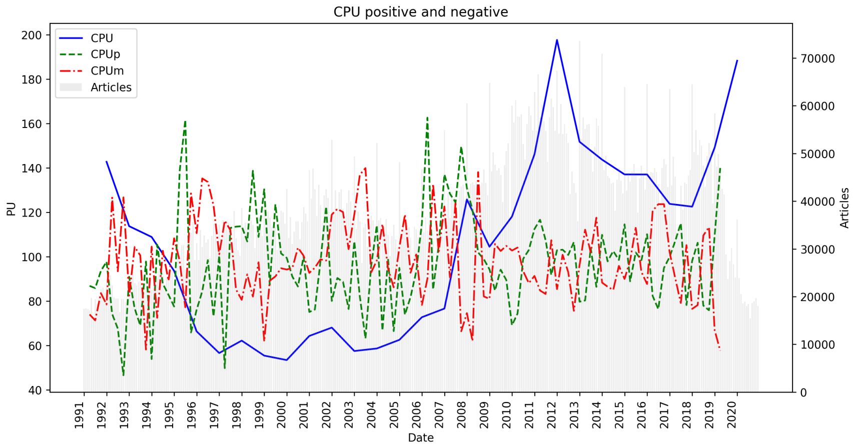
CPU might be pointing in two different directions: at a strengthening or a weakening of future climate stringency, for example suggesting further implementation, or a slowdown in the policymaking process. Previous attempts (Berestycki et al.,

[4] In addition, some news articles refer to specific place-based policies (oil leakages, or polluting plants) which might not necessarily have national relevance. The distinction between local or nationally relevant events is an interesting avenue for further research, but beyond the scope of this paper.

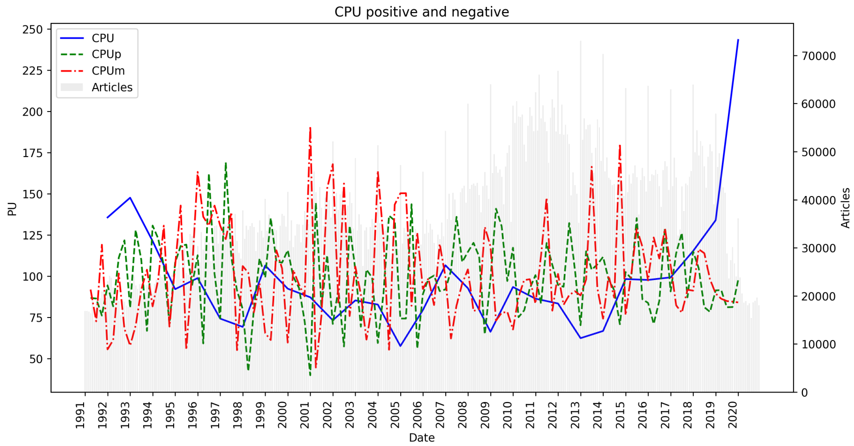
2022; Basaglia et al., 2021) have mapped the direction of CPU, creating sub-indexes for CPU+ and CPU- (increasing or hindering climate policies) for English-speaking countries, based on sets of keywords capturing the direction of uncertainty. Previous measurement exercises in policy uncertainty have employed human labellers for the validation of these indexes, in particular to filter out false positives. I take a different approach, exploiting full-text data. Recent literature has shown the potential of Large Language Models (LLMs) for accurate annotation of textual data, even over-performing crowd labeling (Gilardi et al., 2023). LLMs, in fact, are opening the possibility to lower dramatically the cost of labeling while still achieving human-level accuracy on a variety of different tasks (for a recent review and application, see Wang et al. 2024).

Building on this recent computational social science literature, I prompt the OpenAI API for labeling the universe of CPU news. I make use of the most recent ChatGPT-4o model.[5] I ask three questions during the labeling process. The first serves in filtering out further false positives resulting from the keywords matching: *"Is this news article about climate policy issues?"*. The second two questions are asked to collect information about the policy stance of news articles, and build indicators of positive or negative CPU. The first question is *"Does this piece of news imply a strengthening or weakening of climate policy?"*, and the second *"Are the consequences of this news positive or negative for the environment?"*. The first question is forced as a binary answer, while I leave the possibility, for the latter two questions, to be answered with negative, positive or neutral labels. I perform validation and experimentation of the results of the prompting on a random sample of CPU articles, in similar fashion to the procedure explained in Berestycki et al. (2022) for keywords selection.

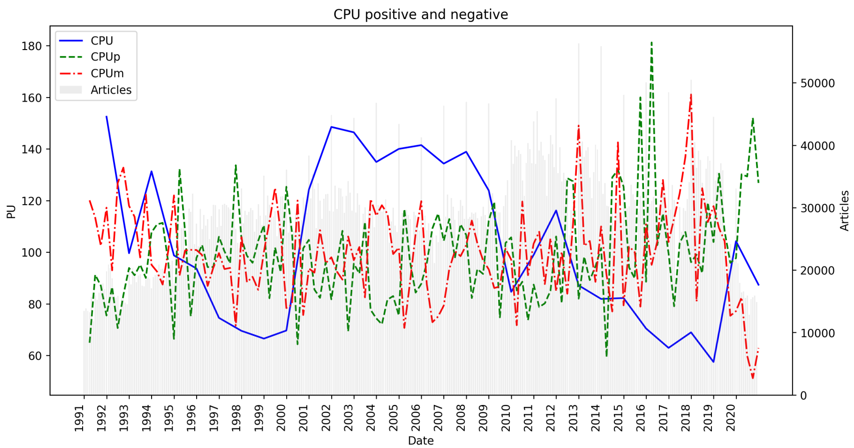
[5] At the time of writing, ChatGPT-4o is the largest LLM model available by number of parameters, estimated at around 1.5 trillion.



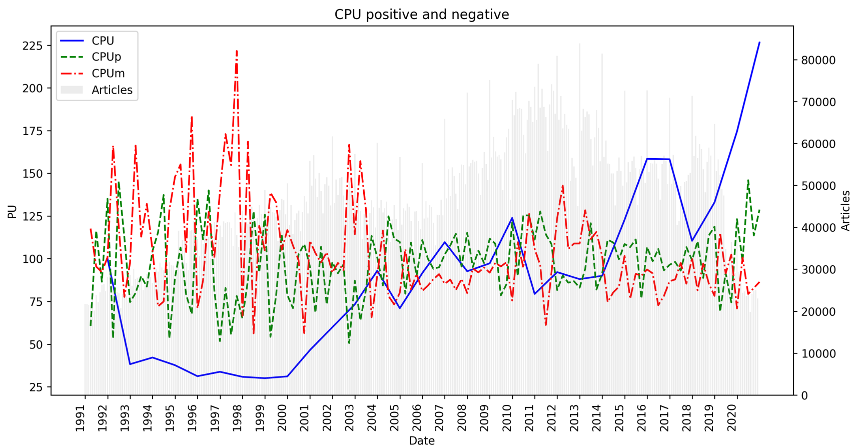
(a) Germany



(b) France



(c) Italy



(d) Spain

Figure 2: Policy uncertainty and policy stance indexes for sample countries

I select as $CPUp$ (CPU plus), articles flagged as true positives during the labeling procedures, and for which any of the two second answers reflect a strengthening of climate policies. Similarly, I flag $CPUm$ (CPU minus), or articles pointing towards a weakening of climate policies or with negative consequences for the environment. Based on this strategy, I derive novel time series, for the four European countries. Rather than dividing the number of articles by the total monthly number of articles in the archive, I divide $CPUp$ and $CPUm$ number of articles by the total number of CPU articles, in order for the series to reflect the relative importance of $CPUp$ or $CPUm$ rather than a general increase of CPU. In Figure 2 I plot the time series for $CPUm$ and $CPUp$, in quarterly moving averages, showing significant variation both over time and across countries.

There are several advantages and disadvantages to the use of LLM technologies for the labeling of articles. Compared to keywords, the labeling process is more of a black-box, while the former is fully reproducible. However, the selection of keywords can be subject to biases and discretionary selections. The multilingual capabilities of LLMs, render this approach well-suited for this dataset, which features four non-English languages. In addition, this flexibility extends in time, mitigating the possible recency bias. Journalism and the use of language changed throughout time, and an approach based solely on frequency might be biased towards more recent policy discussions. Finally, the complexity of syntax-aware methods is particularly important in terms of mapping policy stance. While keyword-dictionaries are based on the simple occurrence of words within documents, LLM architectures are syntax-aware, and better able to capture false positives, handling common issues such as negation.[6]

In the Appendix, I provide a sample of articles' titles matched as $CPUp$ and $CPUm$. Interestingly, while belonging to the domain of climate issues, it is clear that other components are still at play, which might be affecting firms' behaviors differently. Many articles are correctly captured under the correct category: President Bush's government agenda for liberalization, opinion pieces on the risks of environmentalism ($CPUm$) or policy announcements about an increase in kerosene tax ($CPUp$). Other news are of local nature (articles about local smog levels or waste management). In this sense, a promising avenue for further research on the measurement and validation of policy uncertainty measures, could be mixing LLM-labeling methods with unsupervised learning to unpack the universe of news into topics, as proposed in the case of EPU by Larsen (2021).

A number of issues remain open with this approach, and will be further discussed in the limitations section. First, while the results seem promising, false positives and noise still affect in the measurement. A more formal validation of the sub-indexes and the labeling performance remains necessary. The performance of LLMs in comparison with human judgment, in social science applications, requires validation, which is not currently implemented for the sake of this analysis, given the need for human-annotated datasets (Pangakis et al., 2023; Törnberg, 2024). In particular, the prompts

[6]For example, the sentence "innovation subsidies will not slow down the climate transition", would be captured by the terms "slow down" in a dictionary-based approach.

I have created to label the news could be further refined by formalizing a validation method for LLM labeling, given a rather cognitively complex task. In future work, I plan on formalizing and validating the prompt form that more accurately can capture the direction of CPU, limiting its subjectivity and increasing accuracy (Juroš et al., 2024).

Nevertheless, this approach seems particularly promising for social sciences, given the large amount of news-data sources now employed in economics, and its potential to develop teacher-student architectures in machine learning applications. In this architectures, the LLM-generated labels are used as training inputs for training smaller text-models (Pangakis and Wolken, 2024). I apply this architecture to my dataset and test the reproducibility of artificially generated labels with smaller text models. I discuss its potential merits in the case of CPU indices, and provide benchmark evaluation metrics in the Appendix.

Finally, in order to derive a set of robustness indicators, I isolate events (peaks) in both of the CPU series. I build a peak detection algorithm based on the rolling mean of the monthly time series. This algorithm detects peaks based on the deviation of future data points from a rolling mean of the series. I run the peak detection based a six-months moving average, built for each time series, with a threshold of two standard deviations. Thus, a peak is detected if the new data-points exceed two standard deviations from the rolling mean calculated on the past data points. For the implementation of the algorithm I follow the approached proposed in Brakel (2014), where new peaks detected influence the series. In the Appendix Figure 3 I show an example of the peak detection process. All indicators are included in the econometric analysis as yearly moving averages.

3.1.2 Firms and patents dataset

In line with the literature on the economics of innovation, I make use of patent data to proxy the technological efforts of firms. The use of patent data as a proxy for innovation has a long tradition. Despite the numerous criticisms, patent databases represent a valuable source of information for firms' technological efforts, and have been shown to map effectively the knowledge generated by firms, regions and countries.

I use the OECD's REGPAT database (Maraut et al., 2008) for deriving patent-based indicators. Patents widely vary in quality, and can be filed into different jurisdictions at different patent offices. In addition, firms can file several patents to protect the same invention. In order to avoid double-counting patents for the same technology, and to focus on high-quality patents, I make use of the Triadic Patent Families (TPF) database in REGPAT. Within triadic patent families, an invention is filed under the three major patent offices in the world: the European Patent Office (EPO), the United States Patent and Trademark Office (USPTO) and the Japanese Patent office (JPO). Making use of REGPAT, I construct a dataset focusing on patent families rather than single patents, following the approach in Aghion et al. (2016). Patenting inventions is a costly process for firms, and more valuable innovations with promising market perspectives are filed in all three offices. Hence, in this work I focus on high-quality inventions.

Patents can be filed under a large number of technological classes, defined under the International Patent Classification (IPC) and under the Cooperative Patent Classification (CPC). I classify technologies under three categories. In line with the eco-innovation literature, I consider green patents technologies that have potential for mitigation or adaptation of climate change. I match patents based on the methodology recently proposed in Favot et al. (2023), building on previous work (Ghisetti and Quatraro, 2017). I match codes at different digits based on the OECD's ENVTECH classification (Haščič and Migotto, 2015) and the algorithm proposed by Favot et al. (2023) on both IPC and CPC codes for patent families. I expand this search by manually adding codes at higher level from the Y02/Y04S technological classification. The number of TPFs identified as green technologies represent roughly 9% of total patent families (in line with the results in Favot et al. 2023). I provide a detailed summary of the codes employed in Table 13 of the Appendix.[7]

In turn, I identify "dirty" patents families, matching polluting inventions, linked to the the emission of greenhouse gases. I consider as dirty patent families for the production of fossil-fuel, combustion engines, electricity production from non-renewable sources, in addition to steam and gas technologies. Again, I adapt previous work from Aghion et al. (2016); Dechezleprêtre and Sato (2017) for polluting technologies, and expand it with a recent classification of fossil technological codes provided by the International Energy Agency (IEA).[8] Table 15 provides a full description of the technologies considered as dirty. Finally, I create a sub-category of dirty technologies: grey patents. Grey technologies render combustion processes more efficient, and have potential of reducing GHG emissions, while still being polluting technologies. Once again, I follow previous work (Dechezleprêtre and Sato, 2017) and provide a breakdown of grey technological codes in Table 14.

I aggregate total, green, dirty and grey patent families as counts by applicant. Following Aghion et al. (2016) I only consider applicants with consecutive observations. Using information on the name and country of applicants in REGPAT's TPF database, I match firms in the ORBIS (Bureau van Dijk) database, exploiting the name-search engine provided by ORBIS, for the applicants having at least one green or dirty patent over the sample period, and with their address in Germany, France, Spain or Italy. Using balance sheet information from ORBIS, I map firms to their main sectors, and collect information on the year of foundation, and the first available balance sheet. I build an unbalanced panel dataset for firms in the four countries. I consider the beginning of the panel the year of foundation where available, and if not the first year of available balance-sheet information. Where the information is not available, I consider as a starting year the year prior to that of the first patent application recorded in

[7]Where a description of sub-codes' purpose is provided, I match codes at a lower depth than the 4-digits macro group indicated, only considering a subset of those technologies. For a more detailed description of the codes employed, please refer to Favot et al. (2023); Haščič and Migotto (2015) and the most recent OECD's ENVTECH search strategy.

[8]I adapted to REGPAT the search strategies for fossil-fuel patents available online from the IEA. For full details, see:https://gitlab.com/iea$_{d}ds_{p}ublic/iea_{p}atstat/$

REGPAT.[9] I consider as the year of invention of each patent family the earliest filing among the patents belonging to that family.

Table 3: Sample of Firms by country

Country	**Firms**	**Patents**	**Green**	**Dirty**	**Grey**
Germany	2780	163089	13982	12707	3619
Spain	192	4000	327	200	27
France	1112	63740	5142	5285	527
Italy	716	17506	1310	1417	262
Total	4800	248335	20761	19609	4435

Table 3 describes the number of firms and patents by country, breaking down patent counts for each technological category. In order to control for the path-dependency of the innovation process, again borrowing from Aghion et al. (2016) I construct several variables for stocks of previous inventions in dirty and clean technologies, and for geographical spillovers of knowledge available to the focal firm, as detailed in the next Section. In addition, to build control variables, I collect data from Eurostat and the OECD to construct sectoral measures of emissions intensity, following Berestycki et al. (2022). I use data on emissions intensities based on environmentally-extended input-output tables. Emissions intensities are defined as GHG emissions embodied in final demand, from Yamano and Guilhoto (2020), normalized by unit of output.[10] Finally, I collect country-level data for the OECD's Environmental Policy Stringency Index (Botta and Koźluk, 2014).

3.2 Methodology

To investigate the relationship between CPU and the innovation dynamics in firms, I closely follow and adapt the model proposed by Aghion et al. (2016). I test the hypothesis making use of two symmetric models. In the first model, I regress the count of green patent, by year and firm, against Climate Policy Uncertainty and a set of controls:

$$\begin{aligned} PAT_{i,t} = \exp(\alpha + \beta_2 CPU_{i,t-3} + \beta_3 EPS_{c,t-1} * GHG_{c,s,t-1} \\ + \beta_4 K_{i,t-1}) + \eta_i + \tau_{c,t} + \psi_{s,t} + \epsilon_{i,t} \end{aligned} \tag{1}$$

where:

- $K_{i,t}$ is the firm's past patent stock;

[9] Additionally, I correct for discrepancies between firm and patent data. I consider as the starting year the one prior to the first application filing, for firms in which balance sheet information is available, or recorded only after the filing of the first patent.

[10] I consider a number of alternatives for sector emissions, including CO2 intensity per unit of value added, available from the IEA.

- $\tau_{c,t}$ are country by year fixed effects;
- $\psi_{s,t}$ are sector by year fixed effects;
- $\eta_{i,t}$ are firm fixed effects;
- $\epsilon_{i,t}$ is the idiosyncratic error term.

Patent flows are built for three sets of technologies: green, dirty and grey, as count variables by firm and year. I construct the exposure to CPU for the focal firm i, similarly to how Aghion et al. (2016) construct their variable for fuel prices, reflecting the importance of country c for firm i in terms of exposure to policy uncertainty. Firms, in fact, are not subject to Climate Policy Uncertainty deriving from only the country in which they are headquartered, but CPU is weighted by the average share of inventors that the firm has in that country. Inventors' shares are built using REGPAT's database, and each CPU measure is in turn constructed as:

$$CPU_{i,t} = \sum_{c \in C} w_{i,c} * CPU_{c,t} \tag{2}$$

Where $w_{i,c}$ is a time-invariant, firm-specific weight, where $w_{i,c}$ is the (average) share of inventors of firm i in country c, over the period of observation for the firm. Inventors are drawn from the patents' database, and they are assigned to both a country and a firm, based on available information on inventor's location. While inventors could have moved throughout time, I build this indicator as the average share of inventors that each firm has in each country c . I construct identical measures for the directional indicators of positive CPU (variable $CPUp$) and negative CPU (variable $CPUm$). In the main specifications testing for the hypothesis, I include both variables for CPU direction as regressors. In equation 1, I include country by year fixed effects in order to control for macroeconomic conditions and business cycle dynamics that might be correlated with the dependent variables. While country-level policy stringency in environmental regulation should be captured by the country-year fixed effects, I also include an additional control for EPS, interacted with sectoral GHG emissions' intensity, following Berestycki et al. (2022).[11]. Additionally, I include sector by year fixed effects in order to capture sectoral trends that might be correlated with patenting patterns. $K_{i,t-1}$ is the total patent stock of firms, controlling for the size of its innovation portfolio.

Another two symmetric equations are estimated: one for the amount of dirty patents, and one for subset of dirty patents considered grey technologies, as detailed in the previous Section. I time-lag uncertainty and control variables to reflect delayed response, as well as to help mitigate contemporaneous feedback effects. Given the slowdown in investments at a one year lag found in previous exercises, I assume that at least three years should be necessary for the effects to translate onto the outcomes of the innovation process, i.e. on patent applications. I run robustness checks for different lag structures in the robustness Section. In addition, I also lag previous patent stocks, and other controls of one year, reflecting again the path-dependency of the innovation

[11] In robustness checks, I test a different versions of this control, building it similarly to Equation 2

process. Patent stocks are constructed following an inventory rule with depreciation rate r of 20%:

$$K_{i,t} = (1 - r)K_{i,t-1} + PAT_{i,t} \tag{3}$$

The total patent stock of a firm, however, does not allow me to distinguish between green and dirty technologies already available to the firm, as well as the geographical spillovers that might affect patenting patterns. In order to explicitly account for the path dependency of the innovation process, I break $K_{i,t}$ into different components, accounting for both internal and external spillovers, again following closely the approach proposed in Aghion et al. (2016). The stock of knowledge of the firm can be expressed in green and dirty components, both internal and external to the firm:

$$K_{i,t} = GreenStock_{i,t} + DirtyStock_{i,t} + SPGreen_{i,t} + SPDirty_{i,t} \tag{4}$$

Where:

- $GreenStock_{i,t}$ is the firm's own green patent stock;
- $DirtyStock_{i,t}$ is the firm's own dirty patent stock;
- $SPGreen_{i,t}$ are country-level green spillovers to firm i in period t;
- $SPdirty_{i,t}$ are country-level dirty spillovers to firm i in period t;

The stocks of green and dirty patents, for firm i, are again constructed using the inventory method, and control for the path-dependency in the innovation process: the probability of patenting in a specific technology depends on the past track-record of technologies patented in that domain. In addition, country-level spillovers to firm i, control for the external factors that can affect the focal firm's patenting: green (dirty) patenting, can be influenced by the availability of similar technologies outside the firm in that same country. Firms can learn from the available knowledge pool in green and dirty patents, which affects the probability of applying for more patents in the following years. The construction is again symmetrical for dirty and green technologies, and similar to the approach used for that of policy uncertainty. For green technologies, the spillovers available to firm i at time t are:

$$SPGreen_{i,t} = \sum_{c} w_{i,c} * SPGreen_{c,t} \tag{5}$$

The spillover pool in country c ($SPGreen_{c,t}$) is defined as the sum of all other firms' patent stocks of green technologies ($KGreen_{j,t}$):

$$SPGreen_{c,t} = \sum_{j \neq i} w_{j,c} KGreen_{j,t} \tag{6}$$

As detailed in Aghion et al. (2016) and discussed in follow-up works (von Schickfus, 2021), the baseline Directed Technical Change model estimated with two-way fixed effects, would be inconsistent under strict exogeneity, due to serial correlation of the different patent stocks constructed. Thus, borrowing from their approach, I implement the Blundell-Griffith-Van Reenen (BGVR) estimator (Blundell et al., 1999), which relies on the pre-sample mean of the dependent variable in order to proxy for individual fixed effects. This approach is well-suited to patent data, and in empirical setups where data for the dependent variable is available for the pre-sample period.

I run a set of symmetric models for green, dirty and grey patent flows using a Maximum-Likelihood Poisson estimators, accounting for the count data nature of the dependent variables. In addition to the pre-sample mean of the dependent variable, I add controls for firm-level variables. First, I control for the share of patents of the firm in its country-sector (variable $ShPatents$), controlling for potential competition effects. In robustness checks, I also control explicitly for the size of the firms. Because of the limited availability of consistent balance sheet information (employment or total assets) before the 2010 period, I build a time-invariant variable, collecting the last available data point for the firms' assets. I build a categorical variable for the size of the firm (variable $FirmSize$) based on the quartiles of the distribution of assets. All right-hand side variables are log-transformed, including the pre-sample mean. Finally, given some firms have no lagged patent stocks for some periods, I follow Aghion et al. (2016) and add three dummy variables if the green or dirty (lagged) stocks are zero, or if both are zero. In Table 8 of the Appendix, I provide descriptive statistics for all the variables created.

4 Results

I first test the baseline models, regressing the counts of green, dirty, and grey patents against the aggregate index for CPU. In Table 4, I present estimation results, where in all specifications I include the dummies for the absence of green or dirty patents in the past stock of the firms, which are always significant and not reported. In even columns, I add to the baseline specification the control for the share of country-sector patents of the firm. Standard errors are always clustered at the firm level.

Table 4: Poisson Regression - Baseline estimates for CPU - Green, Dirty and Grey patents.

	Green		Dirty		Grey	
	(1)	(2)	(3)	(4)	(5)	(6)
CPU	0.0658***	0.0660***	0.0908**	0.0870**	0.1180	0.1159
	(0.0138)	(0.0144)	(0.0359)	(0.0348)	(0.0895)	(0.0905)
GreenStock	0.9670***	0.9648***	0.0878***	0.1007***	0.0384	0.0570
	(0.0174)	(0.0174)	(0.0195)	(0.0202)	(0.0529)	(0.0580)
DirtyStock	0.0398***	0.0536***	0.9418***	0.9437***	0.9952***	1.004***
	(0.0121)	(0.0117)	(0.0169)	(0.0166)	(0.0593)	(0.0555)
SPILLgreen	0.4909**	0.5183***	0.1076	0.1245	0.7554	0.8411
	(0.1940)	(0.1958)	(0.3593)	(0.3784)	(0.9403)	(1.004)
SPILLdirty	-0.3984*	-0.4170**	-0.1201	-0.1140	-1.192	-1.243
	(0.2051)	(0.2055)	(0.3660)	(0.3842)	(0.9333)	(1.001)
pre-sample mean	-0.0206***	-0.0066	-0.0468***	-0.0339***	0.0228	0.0351
	(0.0058)	(0.0061)	(0.0089)	(0.0086)	(0.0456)	(0.0494)
Emit	-0.0480	-0.0466	-0.0853	-0.0884	0.4875*	0.5017*
	(0.0639)	(0.0635)	(0.0681)	(0.0660)	(0.2599)	(0.2581)
EPS*Emit	0.0110	-0.0000	0.0449	0.0309	0.2935	0.2620
	(0.0479)	(0.0475)	(0.0484)	(0.0474)	(0.2373)	(0.2236)
ShPatents		-0.0514***		-0.0640***		-0.0926
		(0.0120)		(0.0192)		(0.0742)
Observations	86,562	86,562	82,082	82,082	59,452	59,452
Pseudo R^2	0.65247	0.65321	0.76757	0.76821	0.81496	0.81546
RMSE	0.81964	0.82541	0.86330	0.87008	0.50891	0.51270
Country*Year FE	✓	✓	✓	✓	✓	✓
Industry*Year FE	✓	✓	✓	✓	✓	✓

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *CPU* captures firm-level exposure to climate policy uncertainty. *GreenStock* and *DirtyStock* are, respectively, the depreciated stocks of own-firm past green or dirty Triadic Patent Families. *SPILLGreen* and *SPILLDirty* are firm-level geographical spillovers to the focal firm in green and dirty technologies. *Emit* are country-sector-year emissions, and *EPS* is the index of Environmental Policy Stringency. *ShPatents* is the share of firm patents within its country-sector-year. All models contain dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

I find a positive relationship between the aggregate index of CPU for both green and dirty technologies, suggesting that both climate-related technologies are sensitive to increases in aggregate uncertainty. The coefficients for green and dirty own stocks of past patents are positive and significant, as expected. Own patent stocks coefficients have greater sizes depending on the technology observed: green past stocks have a higher magnitude for green technologies than for dirty and grey technologies, and in the latter case are also insignificant. On the contrary, own stocks of dirty innovations have a higher magnitude for dirty and grey technologies. The positive sign for own past stocks of opposite nature (dirty stocks for green, and green stocks for dirty) have a smaller coefficient but are still positive, indicating possible between-technology spillovers within firms. External spillovers $SPILLGreen$ are also positively correlated with green patent flows, while $SILLdirty$ are have a negative correlation, whereas they are insignificant for dirty technologies.

As mentioned, the aggregate index does not allow us to distinguish between the different directions of uncertainty-related indexes. Therefore, in Table 5, I test the two complementary hypotheses developed in Section 2, and include as independent variables of interest the indexes for $CPUp$ and $CPUm$, respectively reflecting CPU capturing the positive or negative direction of uncertainty in terms of environmental regulation.

Table 5: Poisson Regression - Baseline estimates for CPU - Positive and negative policy stance.

	Green		Dirty		Grey	
	(1)	(2)	(3)	(4)	(5)	(6)
CPUp	0.6122**	0.6394**	-0.2737	-0.2948	-1.244**	-1.335**
	(0.2510)	(0.2523)	(0.3014)	(0.3120)	(0.6017)	(0.6204)
CPUm	-0.5453**	-0.5718**	0.3654	0.3836	1.354**	1.446**
	(0.2518)	(0.2532)	(0.2959)	(0.3062)	(0.5983)	(0.6123)
GreenStock	0.9671***	0.9649***	0.0877***	0.1007***	0.0378	0.0566
	(0.0174)	(0.0174)	(0.0195)	(0.0202)	(0.0529)	(0.0580)
DirtyStock	0.0397***	0.0536***	0.9418***	0.9438***	0.9957***	1.005***
	(0.0121)	(0.0117)	(0.0168)	(0.0166)	(0.0593)	(0.0555)
SPILLgreen	0.5073***	0.5379***	0.1062	0.1245	0.7106	0.8024
	(0.1944)	(0.1961)	(0.3598)	(0.3797)	(0.9538)	(1.019)
SPILLdirty	-0.4175**	-0.4390**	-0.1173	-0.1117	-1.141	-1.195
	(0.2049)	(0.2052)	(0.3663)	(0.3852)	(0.9472)	(1.019)
pre-sample mean	-0.0207***	-0.0066	-0.0468***	-0.0338***	0.0229	0.0355
	(0.0058)	(0.0061)	(0.0089)	(0.0086)	(0.0456)	(0.0495)
Emit	-0.0477	-0.0464	-0.0838	-0.0868	0.4939*	0.5099**
	(0.0640)	(0.0636)	(0.0680)	(0.0659)	(0.2598)	(0.2579)
EPS*Emit	0.0097	-0.0014	0.0458	0.0319	0.3022	0.2714
	(0.0481)	(0.0477)	(0.0485)	(0.0475)	(0.2387)	(0.2245)
ShPatents		-0.0516***		-0.0643***		-0.0942
		(0.0120)		(0.0192)		(0.0743)
Observations	86,562	86,562	82,082	82,082	59,452	59,452
Pseudo R^2	0.65254	0.65328	0.76759	0.76824	0.81505	0.81557
RMSE	0.81957	0.82535	0.86289	0.86967	0.50835	0.51212
Country*Year FE	✓	✓	✓	✓	✓	✓
Industry*Year FE	✓	✓	✓	✓	✓	✓

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *CPUp* captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. *CPUm* instead captures environmentally-negative policy stance. *GreenStock* and *DirtyStock* are, respectively, the depreciated stocks of own-firm past green or dirty Triadic Patent Families. *SPILLGreen* and *SPILLDirty* are firm-level geographical spillovers to the focal firm in green and dirty technologies. *Emit* are country-sector-year emissions, and *EPS* is the index of Environmental Policy Stringency. *ShPatents* is the share of firm patents within its country-sector-year. All models contain dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

In line with the expectations in hypothesis 1a, column (2) shows a positive relationship between $CPUp$ and green patenting, while turning negative for $CPUm$. Interestingly, while the directions of the signs are in line with hypothesis 1b, the coefficients for the total of dirty technologies are insignificant. However, in columns (5)-(6), I present the same results only considering the subset of dirty patents comprehending grey technologies, which are in this case significant. In line with hypothesis 1b, coefficients suggest that uncertainty due to potential setbacks in climate policy-making, is positively related with more grey patenting, while the opposite happens in the case of green technologies. These results seem to confirm both hypotheses. Unpacking aggregate Climate Policy Uncertainty reveals a symmetric relationship with green and polluting inventions, suggesting that, depending on its stance, CPU is an important factor for directing technological change.

In order to to confirm these results, I test for different measures of $CPUp$ and $CPUm$. In table 6, in odd columns, I include the ratios between the country-level index for positive-leaning uncertainty ($RatioP$) over the general CPU index and its counterpart $RatioM$. In even columns, instead, I calculate the ratio between the number of positive or negative events over the total number of events (variables $PeaksP$ and $PeaksM$) detected with the algorithm described in Section 3.1.1. Again, these results seem to confirm the two hypotheses, suggesting a that there is a significant relationship between the direction of uncertainty in climate policies, and that of the technological efforts undertaken by firms.

Table 6: Poisson Regression - Baseline estimates for CPU stances - alternative measures.

	Green		Dirty		Grey	
	(1)	(2)	(3)	(4)	(5)	(6)
RatioP	0.7808**		-0.4563		-1.330*	
	(0.3084)		(0.3831)		(0.7177)	
RatioM	-0.6551**		0.5410		1.424**	
	(0.3076)		(0.3689)		(0.7043)	
PeaksP		0.1190***		-0.0067		-0.0312
		(0.0432)		(0.0512)		(0.0688)
PeaksM		-0.0041		0.1085*		0.2242**
		(0.0460)		(0.0597)		(0.0967)
Observations	72,040	72,040	67,809	67,809	49,725	49,725
Pseudo R^2	0.65688	0.65693	0.76741	0.76745	0.81231	0.81239
RMSE	0.86039	0.86051	0.92261	0.92246	0.52828	0.52888
Country*Year FE	✓	✓	✓	✓	✓	✓
Industry*Year FE	✓	✓	✓	✓	✓	✓
Full Controls	✓	✓	✓	✓	✓	✓

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *RatioP* is the ratio between the total level of CPU and its sub-index of environmentally-positive stance. *RatioM* is the ratio between CPU and environmentally-negative CPU. *PeaksP* and *PeaksM*, respectively, represent the total number of events of positive or negative stance, over the total number of events detected by the peak-detection algorithm. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. All models contain full controls and dummies for null past stocks of green and (or) dirty patents. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

4.1 Heterogeneity and robustness

In Table 7 I divide the sample into different historical periods, running separate estimations focusing on the historical evolution of this relationship. Interestingly, while for the period 1995-2005 only green technologies seem sensitive to the directions of policy uncertainty, the coefficients for dirty technologies are again relevant and much higher for the 2010-2020 period, with notable differences between grey and dirty technologies. Different historical phases, seem to suggest a high degree of heterogeneity, across time, both for acceleration and deceleration of policies, and for the development of climate-relevant technologies. This heterogeneity could be driven by acceleration and deceleration of specific policies. In the last decade, in light of an increased implementation of climate policies, green and dirty might be perceived as diametric alternatives, and policy-direction has probably been more credible in terms of their support of one of the two.

Additionally, innovations evolved over time, and the technological linkages between alternative technologies might be changing over time. Grey technologies (on average) appeared more sensitive to $CPUp$ and $CPUm$ then general dirty ones, while the significance seem to be driven by dirty ones in the most recent period. While speculative, these results seem promising in analyzing the dynamic effects of CPU, depending on technological maturity, and in terms of the degree of substitutability and complementarity of technologies with opposite environmental effects.

Table 7: Poisson Regression. CPU stances: Historical analysis.

	1995-2005			2000-2010			2005-2015			2010-2020		
	Green	Dirty	Grey	Green	Dirty	Grey	Green	Dirty	Grey	Green	Dirty	Grey
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
CPUp	0.9749**	-0.4172	-0.8527	0.1774	-2.897*	-2.660**	0.2081	0.3255	-2.280**	2.155***	-3.308***	-1.459
	(0.3897)	(0.5638)	(1.130)	(0.6943)	(1.527)	(1.298)	(0.4515)	(0.6777)	(1.048)	(0.8050)	(0.8258)	(1.339)
CPUm	-0.8883**	0.6099	0.9368	-0.1295	3.004*	2.754**	-0.1710	-0.2453	2.339**	-2.121***	3.285***	1.485
	(0.3847)	(0.5533)	(1.112)	(0.6925)	(1.537)	(1.349)	(0.4587)	(0.6781)	(1.087)	(0.8076)	(0.8196)	(1.336)
Num. Firms	3952	3952	3952	4011	4011	4011	3908	3908	3908	3678	3678	3678
Observations	23,107	22,800	18,207	24,069	23,370	18,016	23,594	22,150	16,300	21,560	18,728	11,971
Pseudo R^2	0.60821	0.80119	0.86894	0.67785	0.76176	0.79337	0.68525	0.75026	0.71657	0.64898	0.68547	0.62463
RMSE	0.72646	1.2568	0.67817	1.0306	0.86549	0.39836	1.0658	0.70783	0.37594	0.71811	0.53636	0.28074
Country*Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry*Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Full Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. $CPUp$ captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. $CPUm$ instead captures environmentally-negative policy stance. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

In the Appendix, I report several robustness checks. In Table 9, I run the same baseline estimates for $CPUp$ and $CPUm$, by weighting regressions by the average stock of patents of the firm. The coefficients for $CPUm$ and dirty patents are of higher magnitude, and more significant if compared with Table 5, suggesting that $CPUm$, pointing a negative direction, might stimulate overall polluting patenting. Weighting also confirms, at the 10% significance level the correlation between $CPUp$ and green patents. Grey patents remain highly sensitive to both directions of CPU, with coefficients of larger magnitudes.

In Table 11 I add two controls. First, I include a control for the measure of Economic Policy Uncertainty (Baker et al., 2016), calculated analogously to that of CPU. In addition, I also include the categorical variable for quartiles of firm size. The direction of the effects is consistent with previous results, with $CPUp$ and $CPUm$ having opposite effects on green and dirty technologies, and again the latter are driven by grey patents. In Table 12 I build a control for Environmental Policy Stringency similar to that I build for my measure of CPU, again substantially confirming the relevance of the CPU sub-indexes for DTC.

In Table 10 I run a leave-one out exercise, excluding one country at the time from the sample. Interestingly, it seems that Germany is driving the significance in results, as it is possible to see in the last three columns. While the numerosity of firms left in the sample could be an issue, this result is very suggestive on the underlying geographical heterogeneity of CPU. A promising direction in policy uncertainty research, is the investigation of cross-country linkages and spillovers of policy uncertainty (Balli et al., 2017; Abakah et al., 2021). One possible reason for this result could be the relative weight that Germany has in both European climate policy-making, and as a power-house for the production of green technologies. Furthermore, the integration of value chains across countries (and between technologies) might also be a factor at play, and the role of technological linkages between products and industries should be further explored.

More puzzling, instead, are the dynamics on the timing of these associations. Figure 4 in the Appendix plots the coefficients and standard errors of $CPUp$ and $CPUm$ tested at different time lags. In the case of green patents, the only significant lag is at 3 years, and the effects disappears in the longer term for all dependent variables. The significance across technologies at t-3, could be confirming the idea that the innovation output takes time to react to a rise in uncertainty. Interestingly, however, short-term correlations with grey inventions seem to be in the same direction as that of green technologies. This refers again to the nature of different technologies as substitutes or complements, revealing a potential dynamic complementarity between grey and green patents. Firms might be adopting a strategic behavior in reducing emissions of their products in the short term, while switching to alternative green technologies in the longer term.

However, further research is needed in this direction to account for the high-degree of volatility in patenting activity. Grey patents only represent a small fraction of total patenting activity, and a macro-level analysis modeling more precisely sectoral dynamics could help clarify this evidence. Moreover, a source of noise could also be the use of the earliest filing applications for patents. Filing years of technologies are an approximation of the timing of innovation activities, but are also the byproduct of legal proceedings, and crucially depend on invention quality. In this sense, keeping in mind that these correlations regard high-end innovation, for which arguably the cost in R&D is higher, looking at the whole spectrum of patent quality could reveal a different pattern of strategic behavior relevant for directed technical change. Evidence on uncertainty and the qualitative features of innovation could render important insights (Bhattacharya et al., 2017). In this framework, the interplay between the business-cycle features of uncertainty and heterogeneity in innovation development could shed further light on the forces directing clean and dirty technological change (Manso et al., 2023).

4.2 Limitations and further research

A number of other limitations apply to this study. First, while being a promising research avenue, linking innovation activities proxied by patents and uncertainty measures, suffers from a discrepancy in frequencies of the data employed. Patent filing dates are relevant at the yearly level, but, as shown in Section 3.1.1, much variation in CPU indexes is lost by aggregating at the yearly level. This loss of information about higher-frequency dynamics limits the understanding of short and long-term behaviors of firms. Other studies employing quarterly measures for investments, relying on firms' reporting, exploit this variation, losing on the technological heterogeneity by considering aggregate investments. Future research could exploit higher-frequency measures, for example in survey data, to bring further evidence on the time-dynamics of green vs brown technologies. As mentioned, these dynamics might also be explored in the light of the linkages between low and high-carbon inventions. As these technologies present spillovers and path-dependencies, there might be supply-chain related factors at play, which policymakers should be considering.

The geographical coverage of this analysis is mainly driven by the availability of adequate sources of news data. The external validity of this study, therefore, warrants a more thorough analysis, especially considering the peculiar evolution of news markets in each specific country.

Furthermore, while the fixed effects strategy employed in this paper should capture relevant confounders, more econometric work is necessary in order to asses the causality of the relationships uncovered. First, the use of pre-sample means and the BGVR estimator could not be fully capturing firm fixed effects. Similar approaches, both in terms of control function estimations and structural modeling could be suitable to address these issues. Second, the nature of firms is increasingly global. Better data is necessary in order to disentangle their geographical presence, both for emissions (which can have strongly local components), and for the geography of their intellectual property protection. Additionally, a number of political economy concerns could be biasing these results, specifically in terms of reverse causality. Bigger firms and more dominant technological actors, arguably have a much higher potential for influencing government action via lobbying efforts, or by setting an anti-environmental agenda in the media. While these concerns are mitigated by the fact that climate policies are often discussed in an inter-governmental setup, and that the consensus for action is confirmed by international treaties, lobbying could cause increases in CPU, or in one of its sub-components.

Nevertheless, this exercise uncovers another potential avenue for future research. The lack of sectoral variability in CPU indexes (and most of the available EPU ones) could be an important gap to fill. Sub-indexes derived for $CPUp$ and $CPUm$, and are blind to sectors and technologies. Arguably, both climate policies and CPU are indeed sector-specific, if not technology-specific, thinking for example to innovation subsidies. Recent papers (Juhász et al., 2022; Evenett et al., 2024) applied text-as-data techniques to the categorization of policy texts, quantifying (green) industrial policy efforts. In this sense, empirical applications based on rich textual data, as in this work, could give nuance to uncertainty measures, adding sectoral or technological dimensions. In the same vein of (Gugler et al., 2024), heterogeneity in CPU linked to specific policy instruments (subsidies, carbon taxes, etc.) could be explored. A promising approach for future applications is the mix of policy-stance with content analysis, breaking down policy-uncertainty into sub-components.

5 Conclusions

In this chapter, I investigate the role of Climate Policy Uncertainty for directing technical change. I build a novel dataset by scraping newspaper archives for four European countries: Germany, France, Italy and Spain. I apply text-as-data techniques to derive sub-measures for policy stance underlying CPU, showing a high-degree of variation between positive-leaning and negative-leaning articles. I bring forward additional empirical evidence by constructing a panel of European firms, and testing relevance of CPU sub-indexes in directing environmentally-sensitive technologies. I employ a model of directed technical change and study patenting activity of firms in both low-carbon and polluting technologies.

I find that CPU pointing towards stronger climate policy implementation is positively associated to the development of green technologies, and instead negatively to polluting ones. On the contrary, the measure of CPU implying setbacks weakening climate policy, shows symmetric results, by favoring dirty innovation and discouraging low carbon inventions. These results suggest that CPU might be affecting firms' expectations about the future value of environmentally-sensitive inventions and the direction of their R&D efforts. I propose a novel approach to identify CPU articles from newspaper data, showing that it could be flexibly applied to different data sources in multilingual contexts, through the use of supervised deep learning architectures.

In line with the extant literature, I find that not only realized climate policy but also uncertainty about the probability of future policies affects firm behaviors (Basaglia et al., 2021; Khalil and Strobel, 2023). These findings suggest a complex and dynamic response, in terms of firms behavior, to differently-leaning uncertainty, and show the relevance of CPU directionality in the belief revision of firms. I add to the previous literature by showing that the direction of this probability has opposite effects on green and dirty innovations. Directing the economy away from a carbon intensive equilibrium to a cleaner growth path is a priority for policymakers, and governments have been experimenting with mission-oriented policy agendas for sustainability (Mazzucato, 2018), and new forms of green industrial policy (Rodrik, 2014). Consensus, clarity and communication surrounding climate and green industrial policies is deemed even more relevant in light of the significant effects on patenting patterns. Legislators can accelerate divestment from fossil technologies and foster green growth by committing to a decisive climate policy agenda. Governments can provide clear signals to the market in support of low-carbon growth, altering firms' expectations and directing innovative efforts towards a cleaner growth path. While preliminary, these results suggest that potential cost of climate policy could be lowered by ensuring coherent market signals to firms. On top of policies, government certainty and commitment to a strong climate agenda could spur virtuous circles, steering economic growth towards a low-carbon future.

References

Abakah, E. J. A., Caporale, G. M., and Gil-Alana, L. A. (2021). Economic policy uncertainty: Persistence and cross-country linkages. *Research in International Business and Finance*, 58:101442.

Acemoglu, D., Aghion, P., Bursztyn, L., and Hemous, D. (2012). The environment and directed technical change. *American economic review*, 102(1):131–166.

Aghion, P., Ahuja, K., Bown, C. P., Cantner, U., Criscuolo, C., Dechezleprêtre, A., Dewatripont, M., Hausmann, R., Lalanne, G., McWilliams, B., et al. (2023). Sparking europe's new industrial revolution-a policy for net zero, growth and resilience. Technical report, Bruegel.

Aghion, P., Dechezleprêtre, A., Hemous, D., Martin, R., and Van Reenen, J. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy*, 124(1):1–51.

Altenburg, T. and Rodrik, D. (2017). Green industrial policy: Accelerating structural change towards wealthy green economies. *Green Industrial Policy.*

Bai, D., Du, L., Xu, Y., and Abbas, S. (2023). Climate policy uncertainty and corporate green innovation: Evidence from chinese a-share listed industrial corporations. *Energy Economics*, 127:107020.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.

Balli, F., Uddin, G. S., Mudassar, H., and Yoon, S.-M. (2017). Cross-country determinants of economic policy uncertainty spillovers. *Economics Letters*, 156:179–183.

Barbieri, N., Marzucchi, A., and Rizzo, U. (2020). Knowledge sources and impacts on subsequent inventions: Do green technologies differ from non-green ones? *Research Policy*, 49(2):103901.

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

Berestycki, C., Carattini, S., Dechezleprêtre, A., and Kruse, T. (2022). Measuring and assessing the effects of climate policy uncertainty.

Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. *The quarterly journal of economics*, 98(1):85–106.

Bettarelli, L., Furceri, D., Pizzuto, P., and Shakoor, N. (2023). *Environmental policies and innovation in renewable energy.* International Monetary Fund.

Bhattacharya, U., Hsu, P.-H., Tian, X., and Xu, Y. (2017). What affects innovation more: policy or policy uncertainty? *Journal of Financial and Quantitative Analysis*, 52(5):1869–1901.

Blundell, R., Griffith, R., and Van Reenen, J. (1999). Market share, market value and innovation in a panel of british manufacturing firms. *The review of economic studies*, 66(3):529–554.

Botta, E. and Koźluk, T. (2014). Measuring environmental policy stringency in oecd countries: A composite index approach.

Bouri, E., Iqbal, N., and Klein, T. (2022). Climate policy uncertainty and the price dynamics of green and brown energy stocks. *Finance Research Letters*, 47:102740.

Brakel, J. v. (2014). Robust peak detection algorithm using z-scores. https://stackoverflow.com/questions/22583391/peak-signal-detection-in-realtime-timeseries-data.

Cascaldi-Garcia, D., Sarisoy, C., Londono, J. M., Sun, B., Datta, D. D., Ferreira, T., Grishchenko, O., Jahan-Parvar, M. R., Loria, F., Ma, S., et al. (2023). What is certain about uncertainty? *Journal of Economic Literature*, 61(2):624–654.

Cherif, R. and Hasanov, F. (2019). *The return of the policy that shall not be named: Principles of industrial policy*. International Monetary Fund.

Cui, X., Wang, C., Sensoy, A., Liao, J., and Xie, X. (2023). Economic policy uncertainty and green innovation: Evidence from china. *Economic Modelling*, 118:106104.

Dechezleprêtre, A. and Hémous, D. (2022). Directed technical change and environmental economics. Technical report, Working Paper.

Dechezleprêtre, A., Martin, R., and Bassi, S. (2019). Climate change policy, innovation and growth. In *Handbook on green growth*, pages 217–239. Edward Elgar Publishing.

Dechezleprêtre, A. and Sato, M. (2017). The impacts of environmental regulations on competitiveness. *Review of environmental economics and policy*.

Dell, M., Carlson, J., Bryan, T., Silcock, E., Arora, A., Shen, Z., D'Amico-Wong, L., Le, Q., Querubin, P., and Heldring, L. (2024). American stories: A large-scale structured text dataset of historical us newspapers. *Advances in Neural Information Processing Systems*, 36.

Devlin, J. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Dixit, A. K. and Pindyck, R. S. (1994). *Investment under uncertainty*. Princeton university press.

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*, 16(2):346–356.

Evenett, S., Jakubik, A., Martín, F., and Ruta, M. (2024). The return of industrial policy in data. *The World Economy*.

Fankhauser, S., Bowen, A., Calel, R., Dechezleprêtre, A., Grover, D., Rydge, J., and Sato, M. (2013). Who will win the green race? in search of environmental competitiveness and innovation. *Global environmental change*, 23(5):902–913.

Favot, M., Vesnic, L., Priore, R., Bincoletto, A., and Morea, F. (2023). Green patents and green codes: How different methodologies lead to different results. *Resources, Conservation & Recycling Advances*, 18:200132.

Feng, G.-F. and Zheng, M. (2022). Economic policy uncertainty and renewable energy innovation: International evidence. *Innovation and Green Development*, 1(2):100010.

Feng, Y. and Ma, X. (2024). The impact of perceived environmental uncertainty on green innovation: evidence from text recognition of chinese listed companies' annual reports. *Environmental Science and Pollution Research*, 31(5):7146–7166.

Fusillo, F. (2023). Green technologies and diversity in the knowledge search and output phases: Evidence from european patents. *Research Policy*, 52(4):104727.

Gavriilidis, K. (2021). Measuring climate policy uncertainty. *Available at SSRN 3847388*.

Ghisetti, C. and Quatraro, F. (2017). Green technologies and environmental productivity: A cross-sectoral analysis of direct and indirect effects in italian regions. *Ecological Economics*, 132:1–13.

Gilardi, F., Alizadeh, M., and Kubli, M. (2023). Chatgpt outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30):e2305016120.

Gugler, K., Szücs, F., and Wiedenhofer, T. (2024). Environmental policies and directed technological change. *Journal of Environmental Economics and Management*, 124:102916.

Haščič, I. and Migotto, M. (2015). Measuring environmental innovation using patent data.

Hémous, D. and Olsen, M. (2021). Directed technical change in labor and environmental economics. *Annual Review of Economics*, 13(1):571–597.

Hoang, K. (2022). How does corporate r&d investment respond to climate policy uncertainty? evidence from heavy emitter firms in the united states. *Corporate Social Responsibility and Environmental Management*, 29(4):936–949.

Hu, Y., Bai, W., Farrukh, M., and Koo, C. K. (2023). How does environmental policy uncertainty influence corporate green investments? *Technological Forecasting and Social Change*, 189:122330.

Huang, W. (2023). Climate policy uncertainty and green innovation. *Economics Letters*, 233:111423.

IEA (2020). Iea flagship report - clean energy innovation. https://www.iea.org/reports/clean-energy-innovation.

Iscen, A., Tolias, G., Avrithis, Y., and Chum, O. (2019). Label propagation for deep semi-supervised learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5070–5079.

Juhász, R., Lane, N., Oehlsen, E., and Pérez, V. C. (2022). The who, what, when, and how of industrial policy: A text-based approach. *What, When, and How of Industrial Policy: A Text-Based Approach (August 15, 2022).*

Juroš, J., Majer, L., and Šnajder, J. (2024). Llms for targeted sentiment in news headlines: Exploring the descriptive-prescriptive dilemma. In *Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 329–343.

Khalil, M. and Strobel, F. (2023). Capital reallocation under climate policy uncertainty.

Kyaw, K. (2022). Effect of policy uncertainty on environmental innovation. *Journal of Cleaner Production*, 363:132645.

Larsen, V. H. (2021). Components of uncertainty. *International Economic Review*, 62(2):769–788.

Li, X., Hu, Z., and Zhang, Q. (2021). Environmental regulation, economic policy uncertainty, and green technology innovation. *Clean Technologies and Environmental Policy*, 23:2975–2988.

Losacker, S., Horbach, J., and Liefner, I. (2023). Geography and the speed of green technology diffusion. *Industry and innovation*, 30(5):531–555.

Lütkenhorst, W., Altenburg, T., Pegels, A., and Vidican, G. (2014). Green industrial policy: Managing transformation under uncertainty. *Deutsches Institut für Entwicklungspolitik Discussion Paper*, 28.

Ma, Y.-R., Liu, Z., Ma, D., Zhai, P., Guo, K., Zhang, D., and Ji, Q. (2023). A news-based climate policy uncertainty index for china. *Scientific Data*, 10(1):881.

Manso, G., Balsmeier, B., and Fleming, L. (2023). Heterogeneous innovation over the business cycle. *Review of Economics and Statistics*, 105(5):1224–1236.

Maraut, S., Dernis, H., Webb, C., Spiezia, V., and Guellec, D. (2008). The oecd regpat database: a presentation.

Mazzucato, M. (2011). The entrepreneurial state. *Soundings*, 49(49):131–142.

Mazzucato, M. (2018). Mission-oriented innovation policies: challenges and opportunities. *Industrial and corporate change*, 27(5):803–815.

McDonald, R. and Siegel, D. (1986). The value of waiting to invest. *The quarterly journal of economics*, 101(4):707–727.

Niu, S., Zhang, J., Luo, R., and Feng, Y. (2023). How does climate policy uncertainty affect green technology innovation at the corporate level? new evidence from china. *Environmental Research*, 237:117003.

Noailly, J., Nowzohour, L., and Van Den Heuvel, M. (2022). Does environmental policy uncertainty hinder investments towards a low-carbon economy? Technical report, National Bureau of Economic Research.

Pangakis, N. and Wolken, S. (2024). Knowledge distillation in automated annotation: Supervised text classification with llm-generated training labels. *arXiv preprint arXiv:2406.17633*.

Pangakis, N., Wolken, S., and Fasching, N. (2023). Automated annotation with generative ai requires validation. *arXiv preprint arXiv:2306.00176*.

Pegels, A. (2014). *Green industrial policy in emerging countries*. Routledge London.

Peng, X.-Y., Zou, X.-Y., Zhao, X.-X., and Chang, C.-P. (2023). How does economic policy uncertainty affect green innovation? *Technological and Economic Development of Economy*, 29(1):114–140.

Pindyck, R. S. (2021). What we know and don't know about climate change, and implications for policy. *Environmental and Energy Policy and the Economy*, 2(1):4–43.

Ren, X., Shi, Y., and Jin, C. (2022a). Climate policy uncertainty and corporate investment: evidence from the chinese energy industry. *Carbon Neutrality*, 1(1):14.

Ren, X., Zhang, X., Yan, C., and Gozgor, G. (2022b). Climate policy uncertainty and firm-level total factor productivity: Evidence from china. *Energy Economics*, 113:106209.

Rodríguez-Pose, A. and Bartalucci, F. (2023). The green transition and its potential territorial discontents. *Cambridge Journal of Regions, Economy and Society*.

Rodrik, D. (2014). Green industrial policy. *Oxford review of economic policy*, 30(3):469–491.

Schmitz, H., Johnson, O., and Altenburg, T. (2013). Rent management–the heart of green industrial policy. *IDS Working Papers*, 2013(418):1–26.

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

Törnberg, P. (2024). Best practices for text annotation with large language models. *arXiv preprint arXiv:2402.05129*.

von Schickfus, M.-T. (2021). Institutional investors, climate policy risk, and directed innovation. Technical report, ifo Working Paper.

Wade, R. H. (2014). Return of industrial policy? In *Economic Policies of the New Thinking in Economics*, pages 76–92. Routledge.

Wang, J. B. (2022). Waiting or acting: The effects of climate policy uncertainty. *Available at SSRN 4028885*.

Wang, M., Wurgler, J., and Zhang, H. (2023). Policy uncertainty reduces green investment. *Available at SSRN 4474726*.

Wang, X., Kim, H., Rahman, S., Mitra, K., and Miao, Z. (2024). Human-llm collaborative annotation through effective verification of llm labels. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pages 1–21.

Xu, Y. and Yang, Z. (2023). Economic policy uncertainty and green innovation based on the viewpoint of resource endowment. *Technology Analysis & Strategic Management*, 35(7):785–798.

Yamano, N. and Guilhoto, J. (2020). Co2 emissions embodied in international trade and domestic final demand: Methodology and results using the oecd inter-country input-output database.

Yu, H. and Chen, H. (2023). Economic policy uncertainty and green innovation: Does corporate social responsibility matter? *Finance Research Letters*, 58:104649.

Appendix

Table 8: Descriptive Statistics

	Count	Mean	Std	Min	Median	Max
CPU	101919	82.76	45.56	0.00	75.44	243.28
CPUp	101919	81.82	31.45	0.00	88.88	139.97
CPUm	101919	81.31	31.43	0.00	89.42	145.27
Green	101919	0.20	1.71	0.00	0.00	114.00
Dirty	101919	0.17	2.72	0.00	0.00	363.00
Grey	101919	0.04	1.53	0.00	0.00	236.00
Green stock	101919	0.96	7.00	0.00	0.00	356.25
Dirty stock	101919	0.87	11.74	0.00	0.00	1075.12
SPILLGreen	101919	4420.78	3876.12	11.46	3458.82	37151.78
SPILLDirty	101919	3026.18	2312.28	9.43	2620.57	17798.43
EPS	101919	2.45	1.51	0.33	2.46	5.17
Emit	94377	0.02	0.18	0.00	0.00	8.22
ShPatents	94377	0.09	0.24	0.00	0.00	1.00

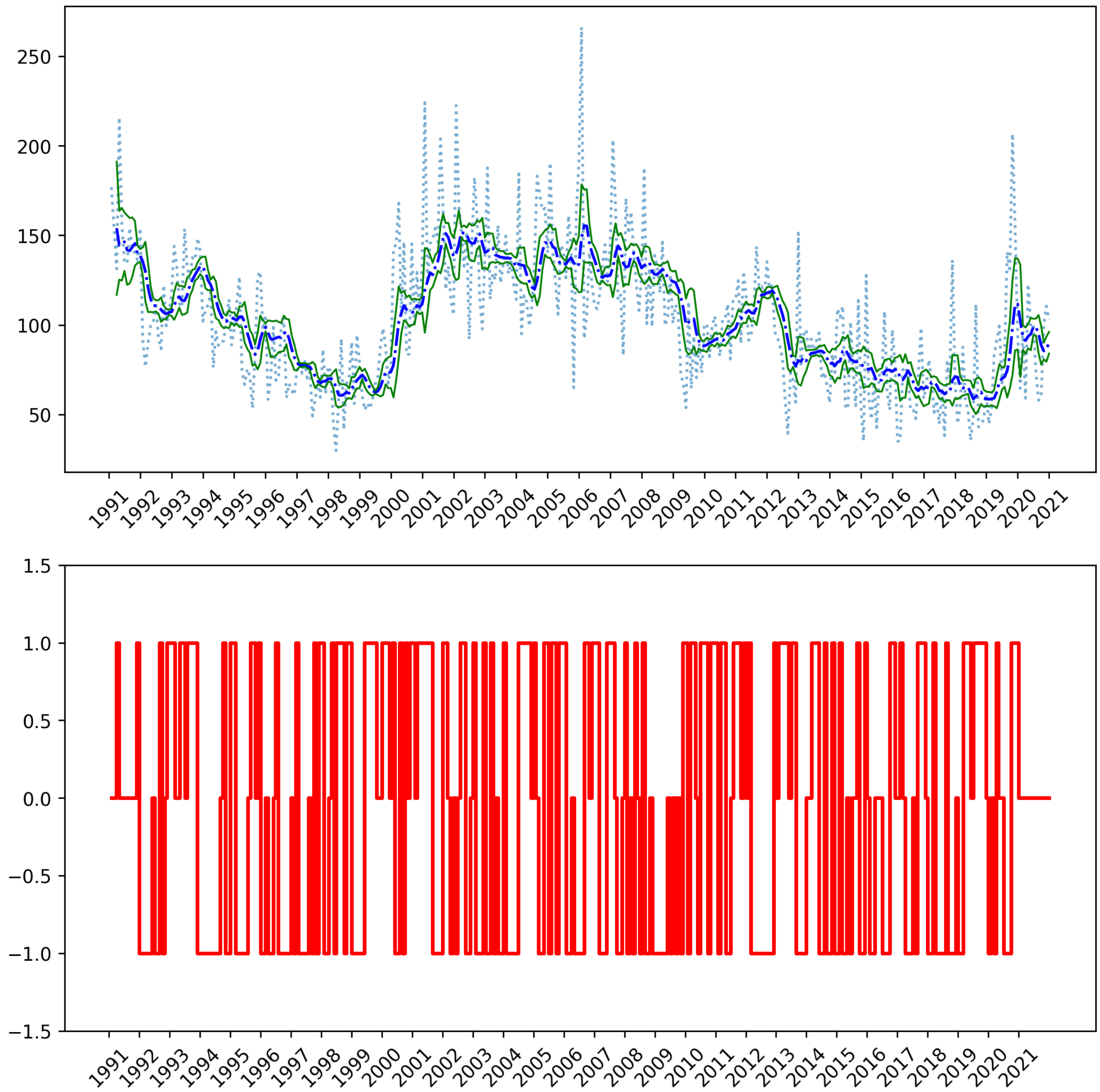


Figure 3: Peak detection algorithm.
The dotted line plots the monthly series of CPU for Italy. In blue, I represent the moving average, and in green the threshold standard deviations for detecting peaks. In the bottom panel, events are flagged as 1 (peaks) or -1 (troughs). I only consider positive deviations (peaks) in the count of events for each time series derived.

Table 9: Weighted Poisson Regression - Baseline estimates for CPU - Positive and negative policy stance.

	Green		Dirty		Grey	
	(1)	(2)	(3)	(4)	(5)	(6)
CPUp	0.6190*	0.6487*	-0.5998	-0.6014	-1.773**	-1.793**
	(0.3597)	(0.3615)	(0.3840)	(0.4053)	(0.8453)	(0.8381)
CPUm	-0.5311	-0.5518	1.078**	1.074**	2.574***	2.603***
	(0.3538)	(0.3545)	(0.4188)	(0.4322)	(0.8620)	(0.8372)
GreenStock	1.034***	1.017***	0.1459***	0.1573***	-0.0278	-0.0256
	(0.0477)	(0.0427)	(0.0322)	(0.0331)	(0.1203)	(0.1203)
DirtyStock	0.0171	0.0338*	0.9404***	0.9401***	1.282***	1.282***
	(0.0225)	(0.0192)	(0.0289)	(0.0260)	(0.1046)	(0.1041)
SPILLgreen	1.331	1.428	1.358**	1.429**	3.299***	3.346***
	(0.9559)	(0.9558)	(0.6264)	(0.6669)	(1.157)	(1.159)
SPILLdirty	-1.389	-1.430	-1.494**	-1.478**	-3.421***	-3.437***
	(1.047)	(1.046)	(0.6465)	(0.6800)	(1.126)	(1.133)
pre-sample mean	-0.0380***	-0.0211	-0.0771***	-0.0647***	-0.1033**	-0.1008*
	(0.0143)	(0.0130)	(0.0163)	(0.0156)	(0.0525)	(0.0558)
Emit	0.0096	0.0045	0.0036	-0.0006	0.8018*	0.7961*
	(0.1049)	(0.1058)	(0.1008)	(0.0984)	(0.4110)	(0.4084)
EPS*Emit	0.0705*	0.0614	0.0819	0.0844	-0.3819*	-0.3807*
	(0.0419)	(0.0412)	(0.0620)	(0.0554)	(0.2197)	(0.2163)
ShPatents		-0.0438*		-0.0582**		-0.0159
		(0.0256)		(0.0254)		(0.0499)
Observations	86,562	86,562	82,082	82,082	59,452	59,452
RMSE	0.77593	0.77660	0.71955	0.71647	0.37788	0.37771
Country*Year FE	✓	✓	✓	✓	✓	✓
Industry*Year FE	✓	✓	✓	✓	✓	✓
Full Controls	✓	✓	✓	✓	✓	✓

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. *CPUp* captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. *CPUm* instead captures environmentally-negative policy stance. *GreenStock* and *DirtyStock* are, respectively, the depreciated stocks of own-firm past green or dirty Triadic Patent Families. *SPILLGreen* and *SPILLDirty* are firm-level geographical spillovers to the focal firm in green and dirty technologies. *Emit* are country-sector-year emissions, and *EPS* is the index of Environmental Policy Stringency. *ShPatents* is the share of firm patents within its country-sector-year. All models contain dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions, weighted by the average number of patents for each firm, over the period of observation. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

Table 10: Poisson Regression. CPU stances: Split sample analysis.

	Excluding France			Excluding Italy			Excluding Spain			Excluding Germany		
	Green	Dirty	Grey	Green	Dirty	Grey	Green	Dirty	Grey	Green	Dirty	Grey
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
CPUp	0.6848**	-0.4028	-1.460	0.4973	-0.8001*	-2.020***	0.7551**	-0.3604	-1.876***	0.5442	-0.3826	-0.3225
	(0.3392)	(0.4092)	(0.9879)	(0.3690)	(0.4502)	(0.7339)	(0.3359)	(0.4096)	(0.6122)	(0.3909)	(0.6181)	(0.7653)
CPUm	-0.6244*	0.4503	1.485	-0.4372	0.8890**	2.140***	-0.6899**	0.4775	1.975***	-0.5086	0.4382	0.3334
	(0.3380)	(0.3960)	(0.9903)	(0.3704)	(0.4458)	(0.7351)	(0.3382)	(0.4119)	(0.6107)	(0.3904)	(0.6020)	(0.7618)
Firms	3229	3229	3229	3572	3572	3572	4007	4007	4007	1705	1705	1705
Observations	55,106	52,403	37,760	61,602	57,937	42,262	69,026	65,217	48,660	26,563	21,544	9,870
Pseudo R^2	0.66901	0.77451	0.83826	0.67090	0.78022	0.82482	0.66006	0.76938	0.81445	0.59669	0.72441	0.62612
RMSE	0.86068	0.84612	0.54918	0.91030	0.95265	0.54967	0.87477	0.93573	0.53189	0.58320	0.80251	0.41412
Country*Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Industry*Year FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Full Controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. $CPUp$ captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. $CPUm$ instead captures environmentally-negative policy stance. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

Table 11: Poisson Regression - Baseline estimates for CPU stances - Additional controls.

	Green		Dirty		Grey	
	(1)	(2)	(3)	(4)	(5)	(6)
CPUp	0.6803**		-0.4285		-1.645**	
	(0.3099)		(0.3878)		(0.6528)	
CPUm	-0.6011*		0.5408		1.924***	
	(0.3151)		(0.3913)		(0.6860)	
PeaksP		0.1208***		-0.0182		-0.0523
		(0.0452)		(0.0563)		(0.0786)
PeaksM		-0.0144		0.0870		0.2436***
		(0.0493)		(0.0616)		(0.0914)
EPU	-0.0211	0.0073	-0.0374	0.0408	-0.2377***	-0.0474
	(0.0384)	(0.0191)	(0.0569)	(0.0382)	(0.0794)	(0.0585)
Observations	72,040	72,040	67,809	67,809	49,725	49,725
Pseudo R^2	0.65696	0.65713	0.76813	0.76811	0.81637	0.81626
RMSE	0.86229	0.86229	0.92212	0.92288	0.50474	0.50629
Country*Year FE	✓	✓	✓	✓	✓	✓
Industry*Year FE	✓	✓	✓	✓	✓	✓
Full Controls	✓	✓	✓	✓	✓	✓
Firm Size Dummy	✓	✓	✓	✓	✓	✓

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. $CPUp$ captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. $CPUm$ instead captures environmentally-negative policy stance. $PeaksP$ and $PeaksM$, respectively, represent the total number of events of positive or negative stance, over the total number of events detected by the peak-detection algorithm. EPU is the firm-level exposure to Economic Policy Uncertainty. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

Table 12: Poisson Regression - Baseline estimates for CPU stances - alternative construction for EPS.

	Green		Dirty		Grey	
	(1)	(2)	(3)	(4)	(5)	(6)
CPUp	0.6972**		-0.4363		-1.518**	
	(0.3096)		(0.3911)		(0.7262)	
CPUm	-0.6125**		0.5286		1.612**	
	(0.3099)		(0.3848)		(0.7278)	
PeaksP		0.1200***		-0.0091		-0.0483
		(0.0432)		(0.0522)		(0.0766)
PeaksM		-0.0005		0.1113*		0.2342**
		(0.0461)		(0.0593)		(0.1058)
EPS	-0.1253**	-0.0665	-0.0895	-0.0267	-0.2624*	-0.2159*
	(0.0515)	(0.0456)	(0.0668)	(0.0657)	(0.1371)	(0.1145)
Observations	72,058	72,058	67,826	67,826	49,725	49,725
Pseudo R^2	0.65697	0.65704	0.76804	0.76793	0.81585	0.81583
RMSE	0.86151	0.86161	0.91886	0.92111	0.50531	0.50689
Country*Year FE	✓	✓	✓	✓	✓	✓
Industry*Year FE	✓	✓	✓	✓	✓	✓
Full Controls	✓	✓	✓	✓	✓	✓
Firm Size Dummy	✓	✓	✓	✓	✓	✓
Cluster S.E.	Firm	Firm	Firm	Firm	Firm	Firm

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. $CPUp$ captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. $CPUm$ instead captures environmentally-negative policy stance. $PeaksP$ and $PeaksM$, respectively, represent the total number of events of positive or negative stance, over the total number of events detected by the peak-detection algorithm. EPS is the Environmental Policy Stringency index, constructed analogously to CPU. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Heteroskedastic-robust standard errors, reported in parentheses, are clustered at the firm level.

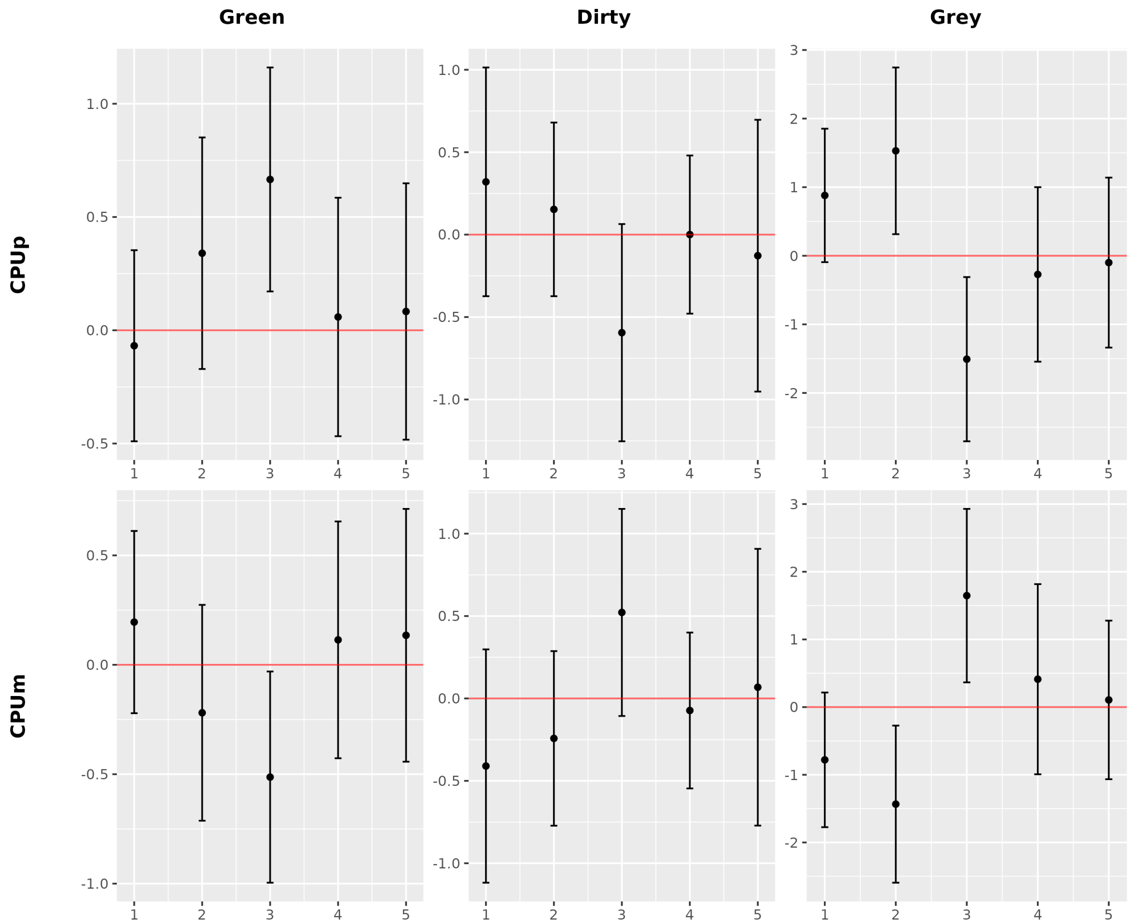


Figure 4: Timing of different lags.

Dependent variables are flows of Triadic Patent Families, identified in their respective technological categories. $CPUp$ captures firm-level exposure to climate policy uncertainty with an environmentally-positive policy stance. $CPUm$ instead captures environmentally-negative policy stance. All models contain full controls and dummy variables in case the past stock of green and (or) dirty patents is 0, which are always significant and not reported. Continuous explanatory variables are log-transformed, applying the inverse hyperbolic sine function. Models are estimated by Maximum-likelihood Poisson regressions. Standard errors at a 5% level are plotted in the Figure.

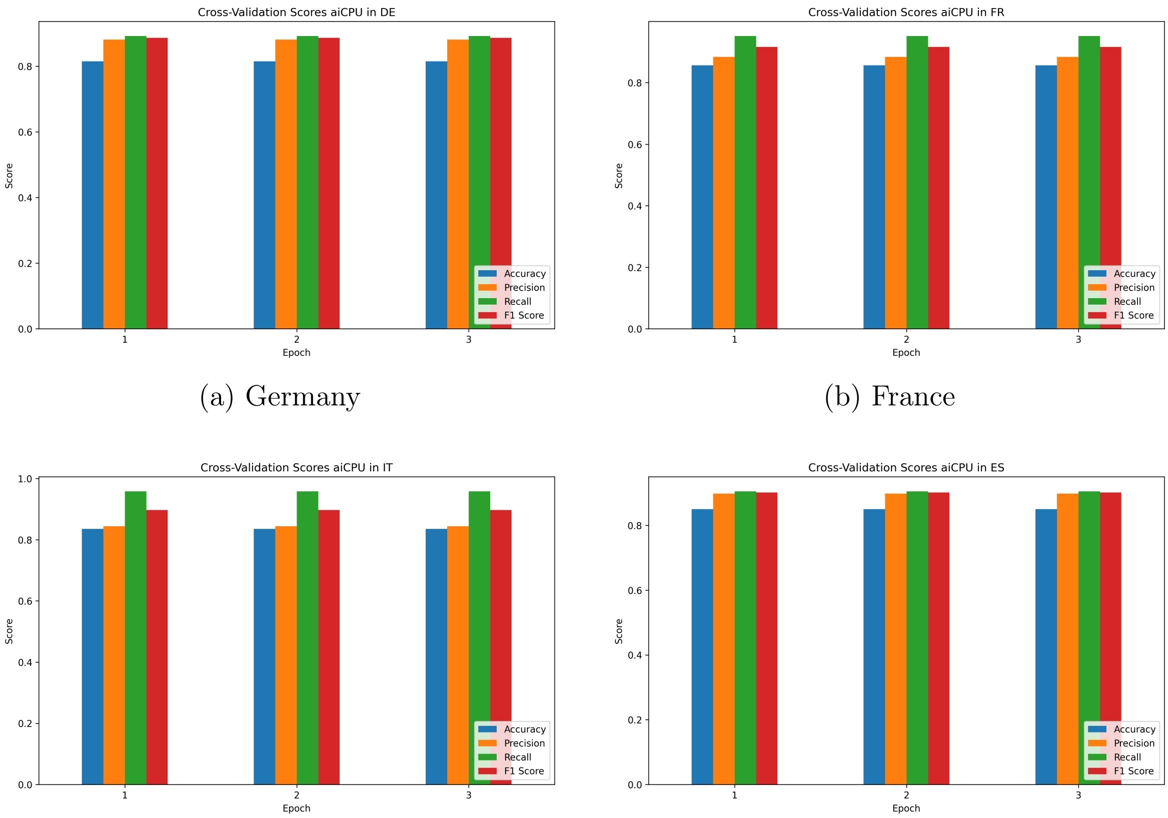
An application of semi-supervised learning in the case of Climate Policy Uncertainty

I provide here an example of architecture for out-of-sample labelling of news data, testing the flexibility of LLM-based methods for labelling of policy uncertainty in newspapers' articles. The objective of this exercise is similar to those of semi-supervised label propagation algorithm (Iscen et al., 2019). In these setups, semi-supervised deep learning exploits a small number of human-curated labelled data, in cases where larger unlabelled data of similar nature are available. The case for this application builds on recent literature using artificial LLM labels as ground truth, and train semi-supervised algorithms on unlabelled data. The algorithm, rather simple in its nature, can be flexibly exploited in the case of policy uncertainty exercises, in which larger amounts of data from a diverse set of news archives can be added.

The architecture is based on recent literature employing LLM labelled dataset, in which a "teacher" algorithm is labelling data subsequently used by another "student" model (Pangakis and Wolken, 2024). This approach, promising in social sciences, can be leveraged to reduce cost of labelling, and potentially achieving human-level quality (Gilardi et al., 2023). After labelling the sample of articles as explained in Section 3.1.1, I extract a random sample of articles amounting to 50% of the dataset, for each country (and therefore each language). In turn, I train a "student" model for each language. Leveraging the superior syntactic properties of tensor-based architectures (in comparison with word-occurrence models), I train Google's Bidirectional Encoder Representation from Transformers (BERT) on two training exercises. First, I train BERT for the prediction of true positives (intended as the LLM-labelled positives) in the sample of Climate Policy Uncertainty articles. Second, I train BERT in a multi-class labelling exercise, based on the direction of climate policy uncertainty labels. The training is performed on the BERT model pre-trained on a large corpus of multilingual data (Devlin, 2018). The relevance of this exercise lies in the possibility to exploit labelled data, costly to obtain, to other news sources within the same language, over a number of different directions. In terms of parameters, BERT is a much smaller model, compared with the latest ChatGPT-4o. Scores for evaluation metrics, across three epochs, are reported in Figure 5, for general CPU, and in Figure 6 for $CPUp$ and $CPUm$. I report scores for accuracy, precision, recall and F1 scores.

The metrics suggest that the prediction of the binary outcome has a quite high predictive power, with F1 scores well above 0.8 in all languages. This result shows that teacher-student architectures are relevant for recognizing news articles about CPU and able to filter out effectively false negatives, also in a multi-lingual context. Trained models, much smaller and cheaper than LLM-based labelling could be applied to new data sources.

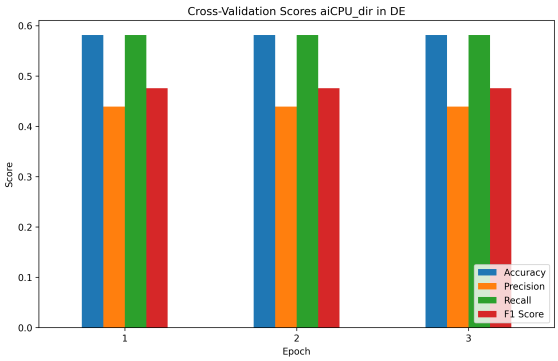
For what concerns the directions of effects, instead, the performance of multi-label is lower. While this attempt could be perfected, human intervention in prompt-tuning, and in the generation of artificial labels might be necessary for achieving a higher out-of-sample performance. However, precision and recall are around 0.6 for all languages, and the series derived from predicted labels (following the methodology explained in Section 3.1.1, correlate at the country level at 0.63 for $CPUp$ and 0.78 for $CPUm$. Overall, these results are very promising for the use of LLM, deep learning and teacher-student architectures in applications related to policy uncertainty, with the potential of making the labelling process cheaper, open source, and flexible and adaptable to novel data sources.



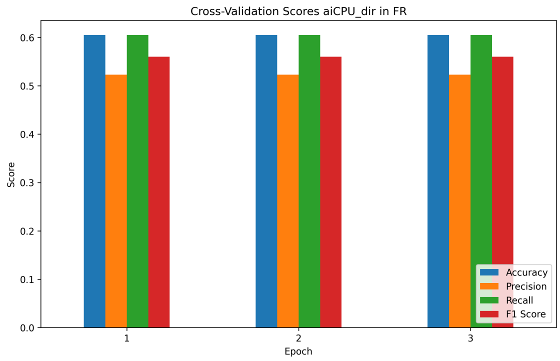
(a) Germany (b) France

(c) Italy (d) Spain

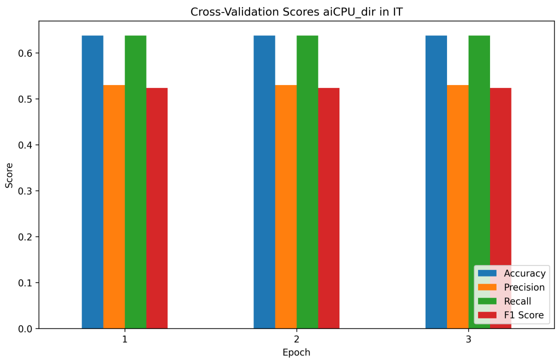
Figure 5: Evaluation scores for three epochs in training - CPU labelling



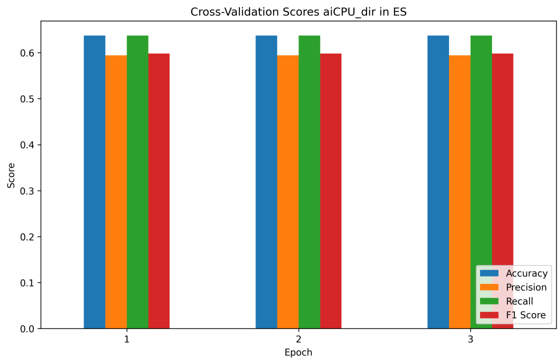
(a) Germany



(b) France



(c) Italy



(d) Spain

Figure 6: Evaluation scores for three epochs in training - CPU directions

Table 13: Technological categories for green patents

Description	Codes	Lower digit purpose
Separation; Purification of air, liquids, or gases	B01D	Used in environmental control technologies such as filtration and water treatment.
Manufacture of iron or steel	C21B	Includes processes that reduce environmental impact in steel manufacturing.
Processing of pig-iron or steel	C21C	Focuses on methods to improve energy efficiency and reduce emissions.
Methods or apparatus for combustion using solid fuel	F23B	Related to cleaner combustion technologies for reduced pollution.
Combustion apparatus using fluid or pulverized fuel	F23C	Involves systems designed for efficient combustion with minimal environmental impact.
Cremation; Incineration of waste	F23G	Waste treatment technologies that minimize emissions.
Removal or treatment of combustion products	F23J	Techniques for controlling and reducing air pollution.
Furnaces; Kilns; Ovens	F27B	Energy-efficient designs and emissions reduction in industrial heating processes.
Chemical or physical processes, catalysts	B01J	Used in environmental applications like pollution control and energy-efficient processes.
Lubricating of internal combustion engines	F01M	Involves technologies to reduce environmental impact from lubricants.
Testing static or dynamic structures, mechanical structures	G01M	Environmental monitoring and testing technologies.
Magnetic or electrostatic separation of solid materials	B03C	Used in recycling and waste processing.
Fuels not derived from petroleum, including biofuels	C10L	Development of cleaner, renewable energy sources.
Exhaust apparatus for combustion engines	F01N	Technologies for reducing vehicular emissions.

Auxiliary equipment for ships, including pollution control devices	B63J	Marine pollution control.
Treatment of water, waste water, or sewage	C02F	Essential for water purification and environmental protection.
Materials for specific applications, including environmental uses	C09K	Involves chemicals for environmental protection like soil conditioners.
Water supply installations	E03C	Technologies improving water efficiency and management.
Sewage disposal	E03F	Waste management and pollution control.
Fertilizers, including those derived from waste	C05F	Involves recycling waste into environmentally friendly fertilizers.
Ships or other waterborne vessels; Equipment for ships	B63B	Marine environmental protection technologies.
Hydraulic engineering; Dams; Harbors	E02B	Involves managing water resources and protecting the environment.
Cleaning streets; Removing snow, ice, or sand	E01H	Environmental management in urban areas.
Collecting or transporting refuse; Containers for refuse	B65F	Waste management technologies.
Animal feeding-stuffs; Non-medical feed additives	A23K	Related to sustainable agriculture and environmental protection.
Footwear	A43B	Involves materials and processes that reduce environmental impact in manufacturing.
Cleaning beaches or sea floor; Other cleaning operations	B03B	Environmental cleanup technologies.
Working of metal powder; Manufacture of articles from metal powder	B22F	Involves sustainable materials and processes in manufacturing.

Preparation or pre-treatment of plastics or other compositions	B29B	Environmental impact reduction in plastic processing.
Presses in general; Pressing	B30B	Includes environmentally friendly pressing methods in manufacturing.
Motor vehicles; Trailers	B62D	Technologies for reducing the environmental impact of vehicles.
Containers for storage or transport of articles or materials	B65D	Involves packaging technologies that reduce environmental waste.
Handling thin or filamentary material	B65H	Includes materials handling in an environmentally friendly way.
Manufacture of glass; Glassware	C03B	Technologies to reduce the environmental impact of glass manufacturing.
Cements; Concrete; Artificial stone	C04B	Focuses on sustainable construction materials.
Working-up of macromolecular substances	C08J	Recycling and environmental impact reduction in polymer processing.
Lubricating compositions	C10M	Development of environmentally friendly lubricants.
Production and refining of metals	C22B	Includes environmental technologies in metallurgy.
Preparation of fibers for spinning; Machines for cotton processing	D01G	Involves technologies for reducing environmental impact in textile manufacturing.
Fibrous raw materials for paper-making	D21B	Environmental impact reduction in the paper industry.
Production of cellulose by removing non-cellulose constituents	D21C	Cleaner technologies in cellulose production.
Pulp compositions; Impregnating materials	D21H	Includes environmentally friendly additives in paper production.
Cables; Conductors; Insulators	H01B	Technologies for energy-efficient and environmentally friendly electrical systems.
Electric discharge tubes; Gas-filled discharge tubes	H01J	Energy-efficient lighting technologies.

Processes or means for direct conversion of chemical energy into electrical energy	H01M	Focus on batteries and fuel cells, including environmentally friendly energy storage.
Sterilization or disinfection techniques; Deodorization	A61L	Environmental impact reduction in medical technology.
Crushing, pulverizing, or disintegrating in general	B02C	Used in recycling and waste processing.
Disposal of solid waste	B09B	Technologies for efficient and environmentally friendly waste disposal.
Cracking hydrocarbon oils; Production of liquid hydrocarbon mixtures	C10G	Cleaner processes in the oil industry.
Reclamation of contaminated soil	B09C	Environmental technologies for soil remediation.
Signaling or calling systems; Preventing, indicating, or extinguishing fires	G08B	Includes environmental monitoring technologies.
Technologies for Adaptation to Climate Change	Y02A	
Climate Change Mitigation Technologies related to Buildings	Y02B	
Capture, Storage, Sequestration or Disposal of Greenhouse Gases	Y02C	
Climate Change Mitigation Technologies in ICT	Y02D	
Reduction of GHG Emissions in Energy Generation	Y02E	

Climate Change Mitigation Technologies in the Production or Processing of Goods	Y02P
Climate Change Mitigation Technologies related to Transportation	Y02T
Climate Change Mitigation Technologies related to Wastewater Treatment	Y02W
Smart grids	Y04S
Fuel Cells	H01M
Wind motors	F03D
Propulsion Of Electrically-Propelled Vehicles	B60L
Tide or wave power plants	E02B9/08
Devices for producing mechanical power from geothermal energy	F03G4
Devices for producing mechanical power from solar energy	F03G6
Ocean thermal energy conversion	F03G7/05
Use of solar heat	F24J2
Production or use of heat, not derived from combustion using geothermal heat	F24J3/08, F26B3/28
Clean Filters	B01D46, B01D50, B01D35, B01D39, B01D41
Water Cleaning	E02B15
Construction waste managment	C04B18

Submerged units incorporating electric generators or motors characterized by using wave or tide energy	F03B13/10-26

Table 14: Technological categories for grey patents

Description	Codes
Idling devices	F02M3/00, F02M3/01, F02M3/02, F02M3/03, F02M3/04, F02M3/05
Injection apparatus in combutions engines	F02M39, F02M41, F02M2041, F02M46, F02M43, F02M45, F02M47, F02M49, F02M51, F02M53, F02M55, F02M57, F02M59, F02M, F02M61, F02M63, F02M65, F02M67, F02M69, F02M71
Adding non-fuel substances to fuel mix	F02M23, F02M25
Electricity control and efficiency	F02D/41, F02B47/06
Combustion technologies with mitigation potential	Y02E20/12, Y02E20/14, Y02E20/16, Y02E20/18, Y02E20/30, Y02E20/32, Y02E20/34

Table 15: Technological categories for dirty patents

Description	Codes
Internal-combustion piston engines	F02B
Controlling combustion engines	F02D
Cylinders, pistons, or casings for combustion engines; arragement of sealings in combustion engines	F02F
Supplying combustion engines with combustiles mixtures or constituents thereof	F02M
Starting of combustion engines	F02N
Ignition (other than compression ignition) for internal-combustion engines	F02P
Oil extraction and refining	C10G
Fuel	C10L1
Separating Plants for Oil-related	B03B9/02, B03D2203/006
Gas-turbine plants	F02C
Production of fuel gases by carburetting air or other gases	C10J
Hydraulic engineering	E02B
Steam engine plants (and similar)	F01K

Steam generation	F22
Combustion apparatus or processes	F23
Furnaces	F27
Heat exchange in general	F28
Lightning	F21H
Conventional and unconventional oil and gas exploratation and extraction	E21B B63B35/4413, B63B2035/442, B63B2035/448, B63B75/00, C09K8, C10L5/04, E02B17/0, E02B2017/003, E02B2017/004, E02B2017/005, E02B2017/006,E21B E02B2017/007, E02B2201,
Exploration and mining	B03B9/0, B03B1, B03D2203/006, B61D11, E21C
Gas conditioning	F25J3/0209, F25J3/0214, F25J3/0615, F25J3/061,
Solid Fuel conditioning	C10F, C10L
Coal to gas processes	C10B47, C10B49, C10B51, C10B53, C10B55, C10B57, C10J1, C10J3
Hydrogen fuel production	C01B3/22, C01B3/3, C01B3/4

Environmental Keywords:
Italian:
((riscaldamento AND (globale OR del pianeta)) OR (emissioni AND NOT (obbligazionarie OR del tesoro)) OR energia OR energetic* OR ambiente OR ambiental* OR ecologic* OR climatic* OR carbonio OR gas serra OR effetto serra OR anidride carbonica OR CO2 OR metano OR CH4 OR inquinament* OR inquinante OR (ossid. AND di zolfo) OR SOx OR diossido di zolfo OR biossido di zolfo OR anidride solforosa OR SO2 OR ossido di azoto OR monossido di azoto OR NOx OR diossido di azoto OR biossido di azoto OR NO2 OR (particelle AND (fini OR solide OR piccole)) OR (particolate AND atmosferic*)) OR polveri sottili OR materiale particolato OR PM10 OR PM2.5 OR ozono OR rinnovabil* OR idroelettric* OR idraulic* OR eolic* OR fotovoltaic* OR emissioni OR biomass* OR (auto OR veicol* AND elettric*) OR ((auto OR motore OR alimentazione) AND ibrid*) OR (solar* AND NOT (crema OR eritema OR sistema OR trattamento OR ustione OR anno)))

French:
((energi* OR énergétiqu* OR environmenta*OR écologique* OR changement climatique OR réchauffement climatique OR climatiq* OR pollution OR pollutan* OR carbone OR gaz à effet de serre OR dioxyde de carbone OR co2 OR ch4 OR méthane OR oxyde de soufre OR so2 OR dioxyde de soufre OR sox OR oxyde d azote OR dyoxyde d azote OR particule fines OR PM2.5 OR PM10 OR ozone OR éolien* OR solair* AND NOT (crème OR système) OR photovoltaïque* OR hydraulique* OR biomasse OR (énergies renouvelables OR énergie renouvelable) OR (voitures OR voiture AND (électriques OR électrique OR hybride*))))

Spanish:
((energ* OR energétic* OR medio ambient* OR ecológic* OR cambio climático OR calentamiento global OR climatic* OR contaminación OR contaminante* OR polución OR carbono OR gases de efecto invernadero OR dióxido de carbono OR CO2 OR metano OR CH4 OR óxido de azufre OR SO2 OR dióxido de azufre OR SOx OR óxido de nitrógeno OR NOx OR dióxido de nitrógeno OR (partículas AND (finas OR en suspensión)) OR PM2.5 OR PM10 OR ozono OR eólic* OR (tecnología* OR panel* OR placa* OR central* AND solar*) OR fotovoltaic* OR (energía AND (hidráulica OR hidroeléctric*)) OR biomasa OR (energías AND (renovables OR verdes OR alternativas OR limpias)) OR (auto* OR coche* AND (eléctrico* OR híbrido*))))

German:
((klima* AND NOT (Geschäftsklima OR politisches OR wirtschaftliches OR Wirtschafts OR Regulierung OR regulatorisches OR Rechts OR rechtliches OR gesellschaftliches OR Gesellschafts)) OR Energiewende OR (Erneuerbare AND Energien AND Gesetz) OR EEG.Einspeisevergütung OR EEG.Umlage OR Klimapolitik OR Energiepolitik OR Umweltpolitik OR Lufteinhaltepolitik OR Luftreinhalteplan OR Umwelt OR ökologisch OR klimawandel OR Erderwärmung OR globale Erwärmung OR Umwelt* OR Energie* OR Kohlenstoff* OR Treibhausgas* OR THG* OR Kohlendioxid* OR CO2* OR Methan* OR CH4* OR Schadstoff* OR Umweltverschmutzung* OR Luftverschmutzung* OR verschmutz* OR schwefeloxid* OR SOx OR Schwefeldioxid* OR

SO2* OR Stickoxid* OR NOx OR Stickstoffdioxid* OR NO2* OR (Partikel* AND (Fein OR Feinpartikel OR Feinstaub)) OR PM2.5 OR PM10* OR Ozon* OR erneuerbar* OR Hydro* OR Windenergie* OR Windpark* OR Windkraftanlage* OR Photovoltaik* OR PV OR Solar* OR Biomasse* OR (Elektrofahrzeug* OR Elektroauto* OR E-auto*) OR (Hybridfahrzeug* OR Hybridauto*)))

Policy Keywords:
Italian:
((politica AND NOT monetaria) OR regolament* OR legislazion* OR legge OR tasse OR canon* OR (standard AND NOT (& OR and OR e OR poors OR poor's)) OR certificat* OR certificazion* OR sussidi OR sussidio OR sovvenzion* OR ETS OR Sistema ES OR feed-in-tariff* OR conto energia OR (scambio AND di quote) OR regime di scambio OR sistema di scambio OR decarbonizzazione OR effetto serra OR cap and trade OR mercato dei diritti di emissione OR (mercato AND (dell OR di AND emission*)) OR (etichett* AND (ambiental* OR ecologic*)) OR eco-etichett* OR eco-label OR normative OR normativa)

French:
((politiq AND NOT monétaire) OR réglementation* OR lois OR loi OR redevance* OR tax* OR impôt* OR norme* OR tarification* OR tarif de rachat OR certificat* OR subvention* OR ETS OR (marché AND d emissions) OR droit* à polluer OR système d échange OR SEQE)

Spanish:
((política AND NOT monetaria) OR regulación* OR ley OR leyes OR impuesto* OR estándar* OR tarifa de alimentación OR certificado* OR subsidio* OR (mercado AND de emision*) OR derecho* OR contaminar OR sistema de comercio OR ETS)

German:
((politik AND NOT geld) OR richtlinie* OR reform* OR regulierung* OR vorschrift* OR gesetz* OR gebühr* OR abgabe* OR maßnahme* OR steuer* OR standard* OR zertifikat* OR subvention* OR preisgestaltung OR emissionshandel OR ETS OR einspeisetarif* OR einspeisevergütung* OR handelssystem* OR cap and trade OR (label OR kennzeichen AND umweltzeichen OR umweltabzeichen) OR umlage)

Uncertainty Keywords:
Italian:
(può OR potrebbe OR probabile OR probabilmente OR possibile OR possibilmente OR potenziale OR potenzialmente OR immaginare OR assumere OR assunzione OR credere OR sostenere OR stimare OR ipotesi OR ipotetico OR speculare OR speculazione OR sospettare OR supporre OR aspettarsi OR dubbio OR dubitare OR dubbioso OR incerto OR incertezza OR sconosciuto OR non familiare OR discutibile OR discutibilmente OR forse OR sembrare OR apparentemente OR improbabile OR nessun indizio OR nessuna prova OR nessuna idea)

French:
(peut OR pourrait OR probable OR probablement OR possible OR possiblement OR potentiel OR potentiellement OR imaginer OR supposer OR supposition OR croire OR prétendre OR estimer OR hypothèse OR hypothétique OR spéculer OR spéculation OR suspecter OR s'attendre à OR doute OR douter OR douteux OR incertain OR incertitude OR inconnu OR non familier OR discutable OR discutablement OR peut-être OR sembler OR apparemment OR improbable OR aucun indice OR aucune preuve OR aucune idée)

Spanish:
(puede OR podría OR probable OR probablemente OR posible OR posiblemente OR potencial OR potencialmente OR imaginar OR asumir OR suposición OR creer OR sostener OR estimar OR hipótesis OR hipotético OR especular OR especulación OR sospechar OR suponer OR esperar OR duda OR dudar OR dudoso OR incierto OR incertidumbre OR desconocido OR no familiar OR discutible OR discutiblemente OR quizás OR parecer OR aparentemente OR improbable OR ningún indicio OR ninguna prueba OR ninguna idea)

German:
(kann OR könnte OR wahrscheinlich OR möglich OR möglicherweise OR potenziell OR potentiell OR vorstellen OR annehmen OR Annahme OR glauben OR behaupten OR schätzen OR Hypothese OR hypothetisch OR spekulieren OR Spekulation OR verdächtigen OR vermuten OR erwarten OR Zweifel OR zweifeln OR zweifelhaft OR unsicher OR Unsicherheit OR unbekannt OR nicht vertraut OR fragwürdig OR fraglich OR vielleicht OR scheinen OR anscheinend OR unwahrscheinlich OR kein Anzeichen OR kein Beweis OR keine Ahnung)

Sample titles of articles flagged as CPU plus:

- L'Europa e la sfida dell'energia
- Il futuro dei Verdi
- Europa è ora di riprendere il cammino
- «Sui rigassificatori intervenga il governo» Scaroni: mi preoccupa il prossimo inverno, ma ci stiamo attrezzando perevitare il peggio
- Bersani «Avanti con le liberalizzazioni di tv ed energia I tagli per risanare il bilancio non si spalmano»
- Rivoluzione energetica contro la crisi
- Rifiuti, Amaie Energia accetta la sfida
- Ue, l'Italia spinge sull'Accordo commerciale per i beni ambientali
- "La infraestructura verde es un símbolo para el nuevo modelo de ciudades"
- La energía solar sale a flote
- «En 10 años ya no podremos invertir el calentamiento»
- «Hay que proteger el paisaje y a los paisanos»
- Blair asegura que ignorar el cambio climático tendrá consecuencias desastrosas
- Lo hemos hecho posible. Ahora tú decides
- «Abrir los mercados no es el nirvana»
- Es hora de tomar en serio el cambio climático
- 2013: le retour en force de l'Europe?
- Une étude de lAPPA précise limpact de la pollution sur la mortalité et la morbidité
- Mis en œuvre avec pragmatisme, un "Green Deal" européen a le potentiel de remodeler l'économie du continent
- L'éolien français manque de souffle
- Nucléaire, éolien... Que proposent Emmanuel Macron et Marine Le Pen en matière dénergie?
- Le contre-modèle américain
- Aérien: la Commission européenne planche sur une taxe kérosène

Sample titles of articles flagged as CPU minus:

- Tutti i rischi dell ambientalismo
- «Meno tasse, meno regole, più sicurezza» Colloquio con Bush, oggi l'insediamento alla Casa Bianca
- Enigma Trump: benvenuti nell'era dell'incertezza
- La fiducia nell'Opec dipende da Mosca
- Economia italiana stabile ma pesa l'incertezza politica europea
- Che cosa succede se gli Stati Uniti abbandonano l'accordo sul clima
- Allarme smog nel giorno dell afa
- La OMC y el futuro del medio ambiente
- Bush propone el mayor aumento del gasto militar desde la era de Reagan
- Un choque de titanes del petróleo en el peor momento posible
- EE UU y China suavizan sus controles medioambientales por la crisis del coronavirus
- Riesgos de catástrofe global
- La política energética de López Obrador provoca incertidumbre en el sector de las renovables
- Écologie et amateurisme
- Comment les Verts ont disparu dune campagne pourtant marquée par lécologie
- Lécologie nest pas morte, cest lécologie politique qui nexiste plus
- Le protocole de Kyoto est moribond, achevons-le !
- Les climato-sceptiques à lassaut du Giec
- Il y a un vrai problème autour de la capacité des Etats en développement à réduire la déforestation
- «La compensation carbone ne doit pas servir à se dédouaner»
- «Les gilets jaunes, symptôme dun peuple qui refuse un monde en perpétuelle accélération»
- Taxe carbone: pourquoi il ne faut plus laugmenter, et même la diminuer!,
- «Greta Thunberg, icône dun écologisme naïf»