Essays on the Role of Policy and Venture Financing in Clean Technology Innovation

Thèse présentée le 28 mars 2022

Collège du management de la technologie
Chaire en politiques d'innovation et de propriété intellectuelle
Programme doctoral en management de la technologie
École polytechnique fédérale de Lausanne

pour l'obtention du grade de Docteur ès Sciences
par

EPFL

Matthias van den Heuvel

Acceptée sur proposition du jury :

Prof Thomas Weber, président du jury
Prof Gaétan de Rassenfosse, directeur de thèse
Prof Ruediger Fahlenbrach, rapporteur
Prof Christopher Tucci, rapporteur
Prof Thomas Astebro, rapporteur

Lausanne, EPFL, 2022

Acknowledgements

I would like to take the opportunity to express my gratitude to Gaétan de Rassenfosse whose kind supervision and valuable support throughout my PhD have been crucial to the successful completion of this thesis. I also want to thank my co-authors Joëlle Noailly and Laura Nowzohour, who have have been tireless companions throughout this journey. Finally, I want to thank David Popp for our collaboration and for his always thoughtful guidance and comments as well as Jacquelyn Pless for her valuable help in the early stages of this project.

I am also grateful for the friendships I have developed over the course of my PhD; in particular with my Gerzensee cohort and with my colleagues at the IIPP and CEMI chairs of the EPFL's CDM. They have brought laughter and joy to my PhD life. Unknowing contributor to this thesis, my family and friends have played a crucial role, allowing me to take a step back and relax when needed.

Finally, my deepest gratitude goes to my parents, without whom I would not have had the intellectual curiosity necessary to embark on this PhD, and to Pauline whose unconditional love, support and enthusiasm have been the fuel of this journey.

Lausanne, April 8, 2022 Matthias van den Heuvel

This thesis has been supported by the Swiss National Science Foundation (SNSF) within the framework of the National Research Programme "Sustainable Economy: resource-friendly, future-oriented, innovative" (NRP 73, Grant-No 407340-172395).

Abstract

This thesis aims to provide novel analyses and data that improve the understanding of the financing of investments in clean technologies. In particular, this thesis explores the role that private and public actors play in supporting young innovative firms.

The first chapter of this thesis exploits original data on about 1000 startups that have participated in Venture Kick, the leading Swiss venture competition. We study the circumstances under which venture competitions' certification and cash prize impact startups' future survival and funding prospects. On the one hand, we observe that the certification only has a significant effect on startups whose quality can be more objectively assessed by experts (e.g., hard sciences startups). On the other hand, we find that the competition's cash prize only extends the runway of startups with low running costs (e.g., internet & mobile). These results highlight industry-specific heterogeneity in the effect of early-stage startup support, which bears implications for the design of entrepreneurial programs.

The second chapter takes a step back and looks into the role of two prominent actors involved in financing new clean technologies: venture capital (VC) and governments. This essay examines the reasons behind the failure of the VC boom in clean energy startups in the early 2010s. We find that lackluster demand for clean energy products and a lower potential for outsized returns make clean energy firms less attractive to VC than startups in ICT or biotech. Conversely, characteristics such as high-capital intensity or long development timeframe seem to be less important. These findings suggest that the most effective way for governments to improve the attractiveness of early-stage clean energy investments is to implement demand-side policies rather than to provide direct financial support. With demand-side policies in place, governments could then try to plug the remaining funding gaps themselves by targeting startups with limited potential for outsized returns.

In the second half of this thesis, the focus shifts onto governments' environmental policies and how they affect private investments in clean technologies. In the third chapter, we develop novel news-based indices of US environmental policy using text-mining techniques. We show that our general environmental policy index and sub-topic indices (e.g., renewables or climate negotiation) meaningfully capture the evolution of environmental policy in the United States. Using these indices, we find that greater salience of environmental policy is associated with increased VC investments in clean technologies and reduced stock returns for high-emissions firms.

In the fourth and last chapter, we study the impact of policy uncertainty on private investment decisions. To do so, we construct a novel index of environmental policy uncertainty (EnvPU). This index reveals that uncertainty is pervasive in the history of environmental policy and is often synchronized with election cycles. We find that increases in our EnvPU index are associated with lower VC investments in cleantech startups and higher volatility in green firms' stocks. The findings from Chapters 3 and 4 suggest that, while the increased salience of environmental policy can boost investments, policymakers should ensure that their actions do not exacerbate policy uncertainty and hinder investments towards a low-carbon economy.

Keywords: economics, clean technologies, entrepreneurial finance, venture capital, environmental policy, innovation, econometrics, text-mining

Résumé

Cette thèse vise à offrir de nouvelles analyses et données améliorant la compréhension du financement des investissements dans les technologies propres. En particulier, cette thèse explore le rôle des acteurs privés et publics dans le soutien aux jeunes entreprises innovantes.

Le premier chapitre de cette thèse exploite des données originales sur environ 1000 startups ayant participé à Venture Kick, la principale *venture competition* en Suisse. Nous étudions les circonstances à travers lesquelles la certification et le *cash prize* des *venture competitions* ont un impact sur les perspectives de survie et de financement ultérieur d'une startup. D'une part, nous observons que la certification n'a d'effet significatif que sur les startups dont la qualité peut être évaluée objectivement par des experts (e.g., en science dures). D'autre part, nous constatons que le *cash prize* ne fait que prolonger la durée de vie des startups à faibles coûts de fonctionnement (e.g., internet & mobile).

Le deuxième chapitre étudie le rôle de deux autres acteurs impliqués dans le financement des nouvelles technologies propres : le capital-risque (CR) et les gouvernements. Cette étude examine les raisons de l'échec du boom du CR dans les énergies propres au début des années 2010. Nous constatons qu'une demande terne et un faible potentiel de rendements démesurés, plutôt que des caractéristiques telles qu'une intensité capitalistique élevée ou un long délai de développement, rendent les entreprises d'énergie propre moins attrayantes pour le CR que les startups dans les TIC ou la biotechnologie. Ces résultats suggèrent que les gouvernements auraient un impact plus important en mettant en œuvre des politiques pour stimuler la demande qu'en finançant directement des startups. Une fois la demande stimulée, les gouvernements pourraient alors essayer de combler eux-mêmes les déficits de financement en ciblant les startups avec un faible potentiel de rendements démesurés.

Dans la seconde moitié de cette thèse, l'accent est mis sur les politiques environnementales et leur impact sur les investissements privés dans les technologies propres. Dans le troisième chapitre nous développons des indices de politique environnementale américaine en utilisant des techniques de *text mining*. Nous montrons que nos indices capturent de manière significative l'évolution de la politique environnementale aux États-Unis. À l'aide de ces indices, nous constatons qu'une prépondérance accrue de la politique environnementale dans les médias est associée à davantage d'investissements en CR dans les technologies propres et à des rendements boursiers réduits pour les entreprises à fortes émissions.

Dans le dernier chapitre, nous étudions l'impact de l'incertitude politique sur les décisions d'investissement privé. Pour ce faire, nous construisons un nouvel indice d'incertitude dans les politiques environnementales (EnvPU). Cet indice révèle que l'incertitude est omniprésente dans l'histoire de la politique environnementale et est souvent synchronisée avec les cycles électoraux. Nous constatons que les hausses de notre indice EnvPU sont associées à une baisse des investissements en CR dans les startups de technologies propres et à une volatilité plus élevée des actions des entreprises vertes. Les conclusions des Chapitres 3 et 4 suggèrent que, bien que l'importance accrue de la politique environnementale dans les médias puisse stimuler les investissements, les décideurs politiques doivent veiller à ce que leurs actions n'exacerbent pas l'incertitude politique et n'entravent pas les investissements vers une économie à faibles émissions de carbone.

Mots-clés : économie, technologies propres, finance entrepreneuriale, capital risque, politique environnementale, innovation, économétrie, text-mining

Contents

Introduction

Motivation

In its 2018 special report, the IPCC reported that to prevent catastrophic damage from climate change, the world must invest between $1.6 and $3.8 trillion in clean energy every year to 2050. However, Bloomberg's New Energy Finance estimates that only $500 billion were invested in energy transition sectors globally in 2020. This means that there needs to be around a fivefold increase in investments in clean energy to meet IPCC targets. To help alleviate this dearth of investment, a greater understanding of the mechanisms behind the financing of clean technologies is crucial.

This thesis focuses mainly on the financing of young firms as they play a critical role in the development of new technologies. The literature has indeed shown that startups contribute disproportionately to economic growth (Audretsch et al., 2006; Haltiwanger et al., 2013) but also to technological progress. Indeed, large incumbent firms tend to improve existing technologies (Acemoglu and Cao, 2010), whereas small firms and new entrants generate more radical innovations (Kamien and Schwartz, 1975). According to the green growth literature, cleantech startups have a similar effect on the development of new clean technologies. Noailly and Smeets (2015) emphasize the role of new entrants, typically small cleantech firms, in driving a technological transition in the energy sector. By contrast, large incumbent firms with a long history of investments in fossil-fuel technologies are subject to path dependencies.

However, it is difficult for new entrants to secure financing for their innovative activities due to factors such as the information asymmetry between entrepreneurs and investors or the lack of assets that can serve as collateral (Gompers and Lerner, 2001; Hall and Lerner, 2010; Ozmel et al., 2013; Robb and Robinson, 2014). This struggle to secure external financing is even more acute for investment in clean technologies. As Howell (2017) puts it, in the absence of a carbon price, "clean technologies [are] at a double disadvantage; they are more difficult to fund both because their positive externalities on climate change are unpriced and because they are more immature, with greater information asymmetry between entrepreneurs and investors". Moreover, in recent years, clean energy startups have found it harder to access venture capital (VC) (Nanda et al., 2015; Gaddy et al., 2017; Sada and Muro, 2017), a crucial source of financing for high-growth but risky startups (Kaplan and Lerner, 2010). This dearth of VC funding makes new sources of early-stage support like venture competitions, accelerators or crowdfunding all the more important (Ewens et al., 2018).

The fact that positive externalities associated with clean technologies are still largely unpriced explains why the development of new clean technologies is often dependent on public support. This dependency on policy support introduces a new type of risk for investors – the risk that policies might experience sudden and unexpected changes – which might deter private investments. In order to study the impact of environmental policies on clean investments and the cost that uncertain and unpredictable policies bear, researchers require meaningful indicators of environmental policy (Brunel and Levinson, 2016). However, high-quality and high-frequency indicators of environmental policy, and a fortiori environmental policy uncertainty, are not yet widely available.

This thesis therefore also aims to produce indicators of environmental policy and environmental policy uncertainty. Moreover, this thesis strives to fill gaps in the literature on early-stage financing of cleantech innovation and in the literature on the impact of policy salience and uncertainty on investments in clean technologies. Findings in these fields could help policymakers better direct and guide sizeable investment efforts towards the cleantech sector.

Overview of the essays and their contribution to the policy debate and the literature

This thesis starts by studying the impact of venture competitions on participating startups' success in Chapter 1 and then takes a step back in Chapter 2 as it explores the role of VC and governments in the financing of cleantech startups. In the second half of this thesis, the focus shifts onto governments' environmental policies. Chapter 3 studies how the salience of environmental policy affects investments in clean technologies, and Chapter 4 measures environmental policy uncertainty and explores whether it hinders investments towards a low-carbon economy. A common thread throughout this thesis is the aspiration to provide data and findings that policymakers can directly use to improve their decision-making.

Chapter 1, in collaboration with Gaétan de Rassenfosse, studies venture competitions, a new type of entrepreneurial program that, like accelerators or incubators, has emerged in the past two decades (Ewens et al., 2018). In this essay, we model and estimate the impact of venture competitions on startup performances using original data on about 1000 startups that have participated in Venture Kick, the leading Swiss venture competition. In particular, we are interested in the effects of the competition's cash prize and certification and why these effects might differ across industries. We find that the certification has a long-lasting effect on the business success of startups whose quality can be more objectively assessed by the competition's judges, notably science-based startups. The information the certification provides accelerates the termination of low-quality startups and improves external funding opportunities for high-quality startups. For startups whose quality is harder to assess, e.g., internet, mobile and software startups, the certification turns out to be too noisy a signal of quality. By contrast, the cash prize does not impact science-based startups' outcomes but significantly extends the runway of startups with lower running costs.

This essay contributes to the nascent literature on venture competitions. At this stage, there is no consensus among scholars as to precisely why these competitions impact startups' outcomes. McKenzie (2017) finds that the cash prize effect matters most, whereas Howell (2020) finds that venture competitions' main benefit is the certification effect. This chapter reconciles these results and argues that, under the right circumstances, both the certification and cash prize can improve startups' prospects.

The first chapter's results bear implications for the design of entrepreneurial programs. The first one is that a competition's certification is informative and improves the allocation of scarce resources. However, for the certification effect to materialize, organizers need to ensure that their judges are able to accurately assess the startups' potential. Only then will entrepreneurs and investors give credence and value to the certification. The second implication is that the cash prizes' role might be narrower than what organizers would hope for, only extending startups' runway without significantly impacting their long-term prospects. Organizers should, therefore, consider how much extra runway their cash prize can buy when designing venture competitions.

As the primary source of early-stage financing, VC has a larger role to play than venture competitions in supporting green innovation. However, a clean energy VC boom in the early 2010s turned to bust as returns proved disappointing. Chapter 2, in collaboration with David Popp, explores the reasons behind the failure of the VC boom in clean energy startups in the early 2010s and the role that governments can play in supporting early-stage clean investments. To do so, we use Crunchbase data on 250,000 companies launched in the United States between 2000 and 2020. We find that weak demand for clean technologies and a low potential for outsized returns can explain why clean energy startups have proven to be poor VC investments. On the other hand, while the capital intensity and long development timeframe of energy startups are often cited as barriers to raising funds for energy firms, we do not find evidence that they are central factors behind the failure of VCs in clean energy. Turning to public investors, we find that they provide useful early and late-stage financing to clean energy startups. However, the ultimate success rate of firms receiving public funding remains small. Therefore, public funding alone is unlikely to bridge the valley of death for clean energy startups.

Thanks to greater data availability today, we can draw more complete lessons from the clean-tech bust than the original assessments (Bumpus and Comello, 2017; Nanda et al., 2015; Gaddy et al., 2017). We find, for example, that investments in digital clean energy startups, which initially seemed to fit the VC model better, have stopped outperforming the rest of the clean energy sector in recent years. Moreover, we offer new evidence of the limited potential for outsized returns of clean energy companies compared to sectors like biotech or information and communication technologies (ICT). Finally, our comprehensive data includes companies in multiple sectors, providing additional context compared to earlier studies focusing solely on public investments in the energy sector (e.g. Howell, 2017; Goldstein et al., 2020).

This chapter provides recommendations to governments on the most efficient policy support to clean energy startups. We argue that governments should first implement demand-side policies that make investing in clean energy more viable rather than investing themselves in startups bound to struggle through the valleys of death. Only then should they turn to financing new firms. That support should focus on firms in cleantech markets where product differentiation is difficult, as such firms are least likely to attract private sector funding.

Chapter 3, in collaboration with Joëlle Noailly and Laura Nowzohour, studies how the salience of environmental policy impacts investments in clean technologies. As a first step, we use supervised machine learning (i.e., text classification) on articles from ten leading US newspapers to construct an index measuring the share of articles about US environmental policy, available monthly over the 1981-2019 period. Then, we apply unsupervised machine learning (i.e., topic modelling) to create 25 sub-indices on policy topics ranging from renewable energy and air pollution to international climate negotiations. We show that our environmental policy indices capture significant policy events in the history of US environmental regulations. Finally, we demonstrate that our environmental policy indices are positively associated with VC investments in cleantech startups and the volume of assets under management of the main clean energy exchange-traded fund.

This chapter contributes to the existing literature in several ways. First, it shows that automated text-mining techniques can help researchers build meaningful newspaper-based indicators. In contrast with developments in macroeconomics, applications of textual analysis in environmental economics remain limited (Dugoua et al., 2021; Baylis, 2020). Second, we make these novel indicators of environmental policy publicly available for other researchers to use (see https://www.financingcleantech.com/envp-index).

As Chapter 3 shows, a supportive policy environment can stimulate private investments in clean technologies. However, if policies are unstable and hard to predict, they can dampen investments; in a recent survey of the European Investment Bank, European firms rank policy uncertainty as the most important barrier to undertaking climate-related investment (European Investment Bank, 2021). Chapter 4 develops a new measure of environmental policy uncertainty, using text mining on the 80,000 environmental policy news articles identified in Chapter 3. This indicator of policy uncertainty provides novel insights into the ebb and flow of uncertainty in US environmental policy. Notably, we find that policy uncertainty is influenced by election cycles and peaks when political power shifts significantly between America's two big parties. We also show further evidence of the negative relationship between policy uncertainty and investments in clean technologies. We find that this relationship is particularly strong for clean energy startups characterized by capital-intensive investments that are difficult to reverse.

The main contribution of this chapter is to propose a new and meaningful indicator of policy uncertainty, a subjective concept difficult to measure quantitatively. We show that our approach based on machine learning algorithms can improve on the current standard approach, popularized by Baker et al. (2016), that relies on the keyword "uncertainty" to identify newspapers articles about policy uncertainty.

In the third chapter, we find that the salience of environmental policy in the media is positively associated with clean technology investments. This finding suggests that news on environmental policy contain relevant information for investors. Therefore, policymakers could leverage their capacity to shape investors' expectations and behavior by communicating audibly about their environmental policy agenda. However, as we show in the final chapter, policymakers can also hinder clean investments if their actions exacerbate policy uncertainty. To limit policy uncertainty, policymakers could, for example, seek to provide stable and predictable policy frameworks decoupled from the volatile politics of election cycles (Aghion et al., 2013).

1 Certification or Cash Prize: The Heterogeneous Effect of Venture Competitions

This chapter is written in collaboration with Gaétan de Rassenfosse.[1]

Abstract:

Despite a rise in scholarly interest in venture competitions, there is still no consensus on exactly why they affect startups' business success. Some scholars argue that venture competitions' main appeal lies in the cash prize they offer. Others claim that what matters most is the certification of startup quality that comes with winning the competition. In this paper, we reconcile these results and argue that, under the right circumstances, both the certification and cash prize can improve startups' prospects. We model and estimate the impact of venture competitions on startup performances using original data on about 1000 startups that have participated in the leading Swiss venture competition. In line with previous findings, we find that venture competitions offer substantial benefits to participating startups. However, these benefits are not distributed equally among participants. In sectors where quality can be more objectively assessed (e.g., hard sciences), startups enjoy long-term value from the certification. The information it provides accelerates the termination of low-quality startups and improves external funding opportunities for high-quality startups. In other sectors (e.g., internet and mobile), the certification proves too noisy a signal of quality. Concerning the cash prize, we find that it only increases the short-term survival of startups with low running costs, with no impact on the more capital-intensive ventures. This finding is evidence that the cash prize's main role is to extend startups' runway. Our results underline industry-specific heterogeneity in startups early-stage support needs, which bears implications for the design of entrepreneurial programs.

[1] The doctoral candidate is the main contributor to this paper.

1.1 Introduction

Venture competitions, events in which startups compete to receive money, mentoring or recognition, have become a key actor in the startup ecosystem. According to Howell (2020), about 15 percent of U.S. startups that secured their first seed or series A financing between 2009 and 2016 have won such a competition. Changes in the investment strategy of venture capitalists (VCs) are reportedly behind the rise of venture competitions. Faced with falling costs of starting a venture, investors have increasingly adopted a "spray and pray" investment approach (Ewens et al., 2018). VCs now provide more limited funding and governance to a greater number of startups—that they are then more likely to abandon.

In this changing environment, venture competitions can provide welcome support to early-stage startups. First, they usually supply a cash prize to the winning startups at a stage where acquiring financing is challenging, thereby mitigating potential financial constraints. Second, they often offer some form of mentoring or networking, presenting fledgling entrepreneurs with opportunities to improve their business. Finally, winning a venture competition can act as a certification of quality, providing valuable information about a startup's probability of success to both founders and investors.

Research on venture competitions is still in its infancy. There is no consensus among scholars as to precisely why these competitions have an impact on startups' business success. McKenzie (2017) finds that it is the cash prize effect that matters most, whereas Howell (2020) finds that venture competition's main benefit is the certification effect. The present paper aims to reconcile these results by exploring factors that drive the certification and cash prize effects.

We start by developing a model of venture competitions' impact on participating startups. Our model yields the following testable predictions. First, certification has a long-lasting effect and is particularly useful to startups in sectors that experts' are adept at screening. Second, because the cash prize's main impact is to extend startups' runway, it has a stronger impact on startups with relatively low running costs, and its effect is limited in time.

To test these predictions, we have obtained original data on 987 startups that participated in Venture Kick, the leading Swiss venture competition, in the period 2008–2017. Switzerland has a vibrant entrepreneurial scene, ranking fourth in the OECD in terms of the amount of venture capital investments per capita. The competition is organized in three eliminatory stages and provides up to $130,000 to winning startups. This set-up is ideal for our purposes as we exploit the multi-stage nature of the competition to tease out the certification and the cash prize effects. Because winning the first stage is very often a startup's first certification of quality but comes with barely any financing, we argue that it primarily captures the certification effect. By contrast, winning the last stage comes with a large cash prize and, therefore, predominantly captures the cash prize effect. Our main identification strategy exploits detailed data on grades given by the judges, which we use to control for differences in startup quality.

Regarding the certification effect, we find that, on average, winning the first stage of the competition increases survival rates two years after the competition by 22 percentage points. It also increases the probability of securing external financing at any point after the competition by 14 percentage points. This is a significant effect considering that the unconditional probability of securing funding after the competition is 26 percent in our sample. However, not all startups benefit equally from the certification. In line with our model's first prediction, we find that the certification effect is particularly strong among startups whose quality can be more objectively assessed by the competition's judges, notably science-based startups (Scott et al., 2020). Winning the first stage increases their survival rate two years after the competition by 41 percentage points. Moreover, this effect is long-lasting, raising survival rates by 35 percentage points six years after the competition. By contrast, startups whose quality is harder to assess, such as internet, mobile & software (ICT) startups, do not seem to benefit from the certification. For these startups, the certification is too noisy a signal of quality.

Turning to the cash prize, we find that, on average, winning the third stage has no causal impact on participating startups' future survival rates or funding outcomes. Once again, though, this result hides heterogeneity across technologies. Indeed, ICT startups with lower running costs are better able to take advantage of the cash prize than the more capital-intensive science-based startups. As a result, receiving the $100,000 prize significantly extends the runway of ICT startups, but has no impact on science-based startups. We find that the cash prize improves the survival rates of ICT startups by 23 percentage points in the first two to three years after the competition, the effect vanishing once the money is exhausted. Finally, the cash prize never seem to affect startups' ability to attract external investors. A sign that the prize money does not fundamentally alter a startup's prospects.

Our results are robust to alternative specifications and identification strategies. Most notably, we implement a sharp regression discontinuity design (RDD) that exploits the fact that startups ranked just above the cutoff receive treatment while those below do not. We also use propensity score matching and errors-in-variables models that allow grades to be observed with some level of imprecision. All approaches point in the same direction.

Another notable finding is that the certification effect seems to work in two ways. First, it acts as a type-revelation mechanism. The feedback provides both winning and losing entrepreneurs with information on their startup's quality that can help them adjust their probability of success. Like Yu (2019) and Howell (2021), we find evidence that losing startups shut down earlier than they would have otherwise. Second, when the judges' evaluation of startups' quality is expected to be more reliable, as in the case of science-based startups, certification provides valuable information on the startups to investors. As a result, winning startups improve their chances of securing external financing after the competition.

It is important to note that our findings only apply to the types of startups that would participate in a venture competition like Venture Kick. We cannot formally test the representativeness of our sample by lack of comprehensive, country-wide data on startups. However, Venture Kick's participants exhibit a value distribution similar to the broader startup population, with a handful of global industry leaders such as Climeworks or Mindmaze and a host of flawed, low-quality startups. Moreover, we find that participants in Venture Kick have rates of exit and post-competition funding similar to the startups that participate in Howell (2020)'s American venture competitions.

The paper contributes to the nascent literature on venture competitions. It offers a broader understanding of the contexts in which a competition's certification and cash prize can benefit startups. We provide evidence that, through their certification, venture competitions can indeed offer valuable information on the startup's quality to both entrepreneurs and outside investors. For this certification to be helpful, however, experts' evaluation needs to be reliable. Regarding the cash prize, we show that its impact depends on how much extra runway it offers. Cash prizes can make a difference when the recipient startups has low running costs (e.g., an ICT startup), the amount given is generous as in Howell (2017) or local costs of living are low as in McKenzie (2017)'s study in Nigeria.

The paper is organized as follows. The next section provides background information on venture competitions. Section 1.3 presents our model and its theoretical predictions. Section 1.4 presents Venture Kick, discusses the data and provides descriptive statistics. Section 1.5 details our preferred empirical strategy. Section 1.6 tests the predictions of our model, investigating Venture Kick's average and industry-specific impact. Section 1.7 confirms the robustness of our results and Section 1.8 concludes.

1.2 Background

The past two decades have seen a sharp decrease in the cost of starting a new high-technology startup or experimenting new ideas.[2] This decrease in costs has sparked change both in the way venture capitalists (VCs) invest and the type of investors that are now able to provide support to early-stage ventures (Nanda and Kropf, 2016; Ewens et al., 2018). VCs have increasingly adopted a "spray and pray" investment approach (Ewens et al., 2018), whereby they make a large number of small investments in early-stage ventures with limited due diligence and governance in the hope that a handful of these ventures will hit a home run. This reduction in VC firms' governance and the fall in the cost of starting a business help explain the emergence of new early-stage financial intermediaries such as venture competitions, accelerators/incubators, angel networks or crowdfunding that provide scalable, lower-cost forms of mentorship (Ewens et al., 2018). In the past decade a literature has emerged to study these new, and increasingly important, financial intermediaries.

Increasingly ubiquitous, accelerators, incubators and university entrepreneurial programs have been documented by many studies (Hochberg, 2016; Gonzalez-Uribe and Leatherbee, 2017; Yu, 2019; Hallen et al., 2020; Eesley and Lee, 2021). Angel investors have also been studied in recent years with Kerr et al. (2011) and Lerner et al. (2018), showing that the participation of

[2] One event that is credited with having sharply reduced the cost of starting a business is the introduction of Amazon's Web Services (AWS) in early 2006, a cloud computing service (Ewens et al., 2018).

angel investors in early-stage ventures increases their survival, employment and follow-on financing. Finally, Mollick (2014) and Hildebrand et al. (2017) analyze the crowdfunding phenomenon spearheaded by platforms like Kickstarter.

Venture competitions have so far received less academic attention, with the notable exceptions of McKenzie (2017) and Howell (2020). Yet, they are an important actor in the startup ecosystem. According to Howell (2020), 14.5 percent of the startups reported on the American data platform CB Insights that secured their first seed or series A financing between 2009 and 2016 won a venture competition. These competitions are events in which a number of startups compete through one or more stages. Their founders pitch their technologies and business models to a panel of judges in the hope of receiving funding, mentoring or certification of quality. The literature has found that venture competitions benefit startups because of either their certification or their cash prize.[3]

Through their certification, the competition's judges provide useful information about the value and potential of participating startups (Scott et al., 2020). This role is essential because information acquisition is crucial in the investment decision of early-stage investors (Bernstein et al., 2017). Gompers et al. (2020) find that selecting the right startup is the most important determinant of VCs' success. Moreover, information asymmetry between investors and entrepreneurs has long been identified as a problem that hinders investment in startups (Amit et al., 1990), and this problem might have become even more acute today. Indeed, the fall in costs has lowered the barriers to entry, making it harder for VCs and other early-stage investors to determine each startup's quality (Ewens et al., 2018). Venture competitions' certification of quality may, therefore, reduce the information asymmetry problem. These competitions also inform the entrepreneurs about their startup's probability of success, guiding their decision to proceed or stop further development efforts. Åstebro and Gerchak (2001) find that the advice to proceed or not with a venture provided by Canada's Inventor's Assessment Program generated higher benefits for society than the costs of running the program. More specifically in the context of venture competitions, Howell (2020) puts forth evidence of a significant certification effect. The author finds that winning a competition seems to primarily serve a certification function, signalling quality to the market and reducing search frictions between VCs and entrepreneurs. Yu (2019) and Howell (2021) find that accelerators and venture competitions provide informative feedback to founders. They argue that this feedback resolves uncertainty around company quality sooner, which accelerates the shutdown of losing ventures.

Turning now to the cash prizes, venture competitions have the potential to alleviate a startup's financial constraints. There is a wealth of evidence that liquidity constraints affect the viability of startups, starting with Holtz-Eakin et al. (1994), and that the growth of most firms is constrained by internal finance (Carpenter and Petersen, 2002). Moreover, Clementi and Hopenhayn (2006) have shown that borrowing constraints have important implications for

[3]Other aspects of venture competitions such as mentoring and networking or the opportunity to learn can benefit participating startups. However, our data are not suited to studying these value-added activities as they are offered to all participants. Hence, we focus on the cash prize and certification effect.

firm growth and survival. By providing a cash prize, venture competitions can typically alleviate some of the winning startups' financial constraints. McKenzie (2017) provides empirical evidence of this effect. Studying a competition in Nigeria that awarded almost $50,000 to its winners, the author finds that winning is helpful mainly because of the large cash prize. Looking at the Shark Tank competition, Smith and Viceisza (2018) also conclude that it is the financing that ultimately matters, and not the signal of quality that winning the show provides.

In order to reconcile these different results, we propose that variations in the competitions' set-ups need to be considered. Everything else equal, it seems logical that a cash prize will be relatively more useful in Nigeria (McKenzie, 2017) than in the United States (Howell, 2020), where startups have access to numerous alternative sources of financing. Moreover, the size of the cash prize relative to the cost of living certainly matters. In the Shark Tank's case, the financing provided is larger than in typical venture competitions, which may help explain why the authors observe a financing effect (Smith and Viceisza, 2018).[4] Overall, these different results suggest that not all participants benefit equally from venture competitions (Lyons and Zhang, 2018). It calls for a more fine-grained understanding of the effect venture competitions on startup success. The following section proposes a model that considers circumstances under which the certification and the cash prize effects will matter to different degrees.

1.3 Theory

1.3.1 Model

We develop a model of the effect of venture competitions from which we draw several testable predictions. We consider three types of agents: entrepreneurs, investors and venture competitions. Entrepreneurs each own a startup i that can be of one of k technologies assigned randomly by nature at time $t = 0$. To operate, startups pay a running cost $C_i(k)$ each period. Entrepreneurs have to decide at the beginning of each period t whether to continue their activity ($y_{i,t}$) or quit. The second agents, investors, own capital and have to decide whether to invest (g_i) in a given startup i.

These decisions are taken with imperfect information because neither entrepreneurs nor investors know the true quality θ_i of startup i. Quality is assigned by nature at $t = 0$ and will influence a startup's future revenues $R(\theta_i, t)$. Revenues also depend on t because they typically increase as a startup transitions from the early stage to its growth stage. Entrepreneurs and investors observe imperfect signals of startup quality, which they use to predict startup quality $\widehat{\theta}(S_{i,k}, Z_{i,t})$. $Z_{i,t}$ is any information about a startup's true quality (e.g., background of the entrepreneurial team) and $S_{i,k}$ is an external quality signal such as the one provided by venture competitions. This predicted quality is then used to estimate expected revenues, $\widehat{R}_{i,t}$. Unlike true quality θ_i, which is fixed, $\widehat{\theta}_{i,t}$ can evolve over time as decision-makers receive more information.

[4] 57 percent of the Shark Tank's contestants receive funding of an average amount of around $145,000.

The third agents, venture competitions, provide participating startups with the following two elements. First, they offer a certification of quality $S_{i,k}$ that provides additional information to both entrepreneurs and investors, allowing them to update $\widehat{\theta}_{i,t}$. Therefore, in this model, certification serves as a type-revelation mechanism rather than as a way to address asymmetry of information. This certification can be either positive (i.e., $S_{i,k} = 1$) or negative (i.e., $S_{i,k} = -1$). Second, the competition gives a financial reward $Prize_{i,t}$ to its winners, whose use is strictly limited to covering the costs of the startup and which is lost if the company stops its activity.

To simplify our infinite-horizon problem formulation, we assume without loss of generality that our agents are fully patient—the discount rate β is thus 1—and we set the risk-free rate at 0. The timing of the model is as follows. In period $t = 0$, nature assigns a quality θ_i and a technology type k to each startup. We assume that the entrepreneurs start with a negligible amount of capital of their own. In period $t = 1$, entrepreneurs receive information about the true quality of their startup via $Z_{i,t}$. They then adjust the estimated quality of their startup $\widehat{\theta}_{i,t}$ and, based on this estimate, decide whether to pursue the business or not. At the beginning of period $t = 2$, some of the startups still alive are randomly selected to participate in a venture competition.[5] Then, from period $t = 2$ onward, entrepreneurs and investors use both $Z_{i,t}$ and $S_{i,k}$ to estimate $\widehat{\theta}_{i,t}$ and make their decisions. The majority of startups will survive in the first periods when $Z_{i,t}$ is small in absolute value. Indeed, we expect entrepreneurs to be optimistic about their probability of success when they have little information about their true quality. This overconfidence of entrepreneurs is well-documented in the literature (Koellinger et al., 2007).

The decision of the entrepreneur

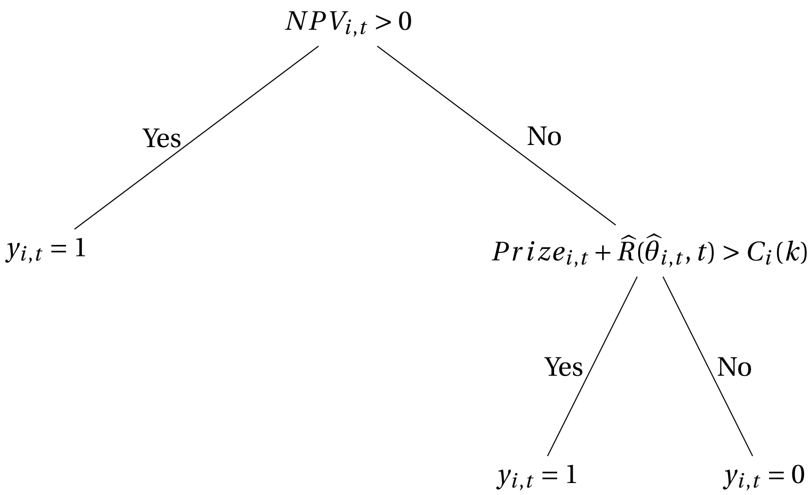
The entrepreneur chooses each period whether or not to continue with their activity—if they shut down in period t, then $y_{i,t+j} = 0 \ \ \forall \ \ j \in [0,\infty]$—based on the expected Net Present Value (NPV) of the venture at time t. The main determinant of the NPV is the startup's estimated quality at time t, $\widehat{\theta}_{i,t}$. Accounting for the fact that a startup might fail and cease to exist, future revenues are weighted by the estimated probability of it being alive n periods in the future, $\widehat{y}_{i,t+n}$. Crucial to both $\widehat{\theta}_{i,t}$ and $\widehat{y}_{i,t+n}$, the information $Z_{i,t}$ is a function of its past value and a random variable $\gamma_{i,t}$ that depends on θ_i. We allow the certification S_i to have a varying impact on $\widehat{\theta}_{i,t}$ depending on how credible and trusted the certification is, which in turn is a function of startup i's technology k. Finally, as periods pass, startups accumulate their own $Capital_{i,t}$ which is the sum of past profits (and does not encompass the non-appropriable cash prize). We assume that outside capital is available to high-quality startups to cover temporary negative cash flows (e.g., from investors). The different variables in the model are

[5] As we are not interested in the decision to participate in a venture competition, we treat participation as random in our model. In reality, participation in a competition is not random, our empirical results therefore have to be interpreted as the effect of venture competitions on the startups that want and are able to participate.

defined as follows:

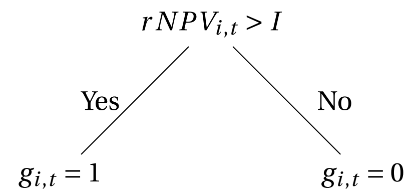
$$
\begin{aligned}
NPV_{i,t} &= \sum_{n=0}^{\infty} \widehat{y}_{i,t+n}(\widehat{R}(\widehat{\theta}_{i,t}, t+n) - C_i(k)) \\
\widehat{\theta}_{i,t} &= f(h_k(S_i), Z_{i,t}) \\
\widehat{y}_{i,t+n} &= f(\widehat{\theta}_{i,t}, t+n) \\
Z_{i,t} &= Z_{i,t-1} + \gamma_t(\theta_i) \\
Prize_{i,t} &= max(Prize_{i,t-1} - C_i(k), 0) \\
Capital_{i,t} &= \sum_{t=1}^{t-1} R_{i,t} - max((t-1)C_i(k) - Prize_{i,2}, 0)
\end{aligned}
$$

At the beginning of each period t, the entrepreneur can be in one of three situations depicted in the decision tree below. To begin with, if entrepreneur i estimates that their startup's NPV is positive they will choose $y_{i,t} = 1$. Alternatively, if an entrepreneur estimates that their project is not viable in the long term (i.e., $NPV < 0$) they have two options. They will choose $y_{i,t} = 1$ if they expect their revenue in period t and what remains, if anything, of the prize to cover the cost of running the business during period t. Pursuing the venture one more period could allow them to increase the entrepreneur's own capital—the prize is lost if the startup stops its activity—or receive new information, $Z_{i,t+1}$, that could make the NPV positive. Finally, they will choose to quit their unprofitable business ($y_{i,t} = 0$) if they cannot cover the costs of running it at time t.



The decision of the investor

The investor chooses once whether to invest in a given startup ($g_{i,t} = 1$ if they decide to invest). Investing means providing I in exchange for a return r on future profits' $NPV_{i,t}$. The $Prize$ does not enter the equation because it cannot be appropriated by the investors and it does not impact the NPV. The investor's decision is simpler than the entrepreneur's: invest if they expects their investment to yield more than its cost and do not invest otherwise.



1.3.2 Predictions

Before drawing some predictions from our model, we need to make explicit two assumptions on which we rely. First, we assume that the certification is an informative signal of quality, that is, $corr(S_{i,k}, \theta_i) > 0$. For certification to be informative, judges need to be able to accurately assess startups' quality. This first assumption is supported by the fact that, as we show in Section 1.5.1, judges' grades in our data are highly predictive of startups' future business success. Second, we allow the judges' ability to identify high-quality startups to differ by technology type. This feature of our model is justified by both our data (see Figure 1.2) and the literature. Indeed, Scott et al. (2020) find that experts can effectively differentiate among early-stage ventures on grounds of quality in the hardware, energy, life sciences, and medical devices sectors but cannot do so for ventures in the consumer products, consumer web and mobile, and enterprise software sectors. A possible explanation for this difference is that the success of science-based startups depends on elements that are more 'objective' compared to consumers products or ICT startups (e.g., hard evidence of a technological advance vs. projection of the number of users of an app). Along this line, science startups in our sample that are alive five years after the competition have 2.6 patents on average compared to only 0.34 for ICT startups. Under this assumption, the competition's certification provides more information on a life science venture than it does on an ICT venture, which implies that $corr(S_{i,k=life\_sciences}, \theta_i) > corr(S_{i,k=ICT}, \theta_i)$.

The certification effect

The certification, $S_{i,k}$, affects how entrepreneurs and investors estimate the startup's probability of success, which improves their decisions (i.e., $\frac{\partial y_{i,t}}{\partial S_{i,k}}, \frac{\partial g_{i,t}}{\partial S_{i,k}} > 0$). On the one hand, startups that lose in the competition (i.e., $S_{i,k}$ goes down) see their estimated NPV going down. The fall in NPV leads them to close down earlier and makes it harder for them to secure funding, thus reducing wastage of resources on low-quality projects. On the other hand, winning startups now have a higher estimated NPV, which leads them to survive for longer and secure more external funding.

Proposition 1: *Certification is particularly useful to startups in sectors that experts' are adept at screening.*

Proposition 1 stems from the fact that a certification will only be useful if it provides trusted and accurate information. However as discussed in Section 1.3.2, the ability of judges to screen startups varies across technology types. We define the sectors where experts can objectively identify startups' quality as $high\_info$ and the sectors where experts struggle to do so as low_info. The certification provides more information about the true quality θ_i of a $high\_info$ startup than of a low_info one. Because of that, we argue that investors and entrepreneurs in the $high\_info$ sectors are more responsive to certification and significantly adjust their estimation of a startup's quality $\widehat{\theta}_{i,t}$ and, consequently, their decisions $y_{i,t}$ and $g_{i,t}$. In other words, because $\frac{\partial\theta_i}{\partial S_{i,k=high\_info}} > \frac{\partial\theta_i}{\partial S_{i,k=low\_info}}$ we predict $\frac{\partial y_{i,t}}{\partial S_{i,k=high\_info}} > \frac{\partial y_{i,t}}{\partial S_{i,k=low\_info}}$.

Proposition 1.a: *The certification has a long-term impact on entrepreneurs and investors.* Additionally, the certification effect is long-lasting. Indeed, the information provided by the certification permanently impacts how investors and entrepreneurs estimate a startup's probability of success and therefore their willingness to invest in or pursue the venture.[6]

The cash prize effect

The cash prize temporarily covers the startups' running costs thereby extending their runway. Receiving a cash prize is particularly useful to startups with negative NPV, allowing entrepreneurs to pursue their activity a few more periods that they could have otherwise. They can use this extra time to gather additional information about to startup's true quality or to increase their capital. Arguably, cash prizes have a less direct impact on funding as investors cannot claim ownership of the prize money and it does not increase the startup's NPV.

Proposition 2: *The cash prize has a stronger impact on startups with relatively low running costs.*
We derive this proposition from the fact that cash prizes impact startups by extending their runway. As a result, the magnitude of the cash prize effect hinges on how much extra runway it buys. This, in turn, depends on the cash prize's size relative to the startups' running costs. Indeed, when a startup develops a capital intensive technology, in other words $\frac{Prize_{i,t}}{C_i(k)}$ is small, a cash prize only has a limited impact. However, the same cash prize will significantly extend the runway of startups with low running costs.

Proposition 2.a: *The cash prize effect is mostly temporary.*
Unlike for certification, we expect the effect of the cash prize effect to be short-lived. The cash prize allows startups of marginal quality ($NPV < 0$) to survive for a little while. But, once the money from the cash prize runs out, entrepreneurs cannot afford to continue to run an unprofitable business and have little choice but to shut-down. Therefore, the positive effect of the cash prize will be temporary for most startups. Only the small number of entrepreneurs

[6]Over time, however, agents receive more and more information about the startup's true quality—i.e., $Z_{i,t}$ grows or falls. The certification's effect might therefore lose some of its prominence as new sources of information become available.

that were able to use the extra runway to discover that their NPV was, in fact, positive will see a long-term positive effect.

1.4 Data

Switzerland and Venture Kick, its leading venture competition, offer an ideal setting to test the predictions of our model. Indeed, Switzerland hosts a thriving entrepreneurial environment with many active early-stage investors. According to the OECD's Entrepreneurship Financing Database, Switzerland is the fourth OECD country with the highest amount of venture capital investments per capita (after the United States, Israel and Canada).[7] This setup is essential because it means that certification will matter and the cash prize will not be the only available source of funding. Additionally, the venture competition we study needs to provide both certification and a cash prize in a way that allows us to estimate their impact separately. These rewards also need to have been provided in a consistent manner throughout the years. Venture Kick satisfies all these criteria. It has maintained a very similar reward structure in terms of certification/publicity and cash prize during the entire study period. Moreover, the judges have graded startups using the exact same guidelines and criteria since the start of the competition in 2007. Finally, because Venture Kick is open to all technology types, we are able to observe significant variability to the parameters that matter in our model: running costs and experts' ability to predict the startup future success.

1.4.1 The Venture Kick competition

Venture Kick is a Swiss venture competition based in Lausanne and Zurich. It has awarded more than $29 million to 675 startups since its launch in 2007 (Venture Kick, 2019). It aims to increase the number of spin-offs at Swiss universities, accelerate the time-to-market and raise the attractiveness of these young companies to professional investors. Venture Kick has become an important actor in the Swiss startup scene. Using data from CrunchBase, we find that 22 percent of all the startups based in Switzerland founded between 2006 and 2019 that completed a round of series A financing received money from Venture Kick. More importantly, seven of the thirteen best-ranked startups in this group, including MindMaze, Switzerland's first unicorn at $1+ billion valuation, are Venture Kick alumni.[8] Venture Kick has, therefore, supported a disproportionate share of Switzerland's most successful startups. This record of identifying successful startups comforts the argument that Venture Kick has the potential to be a credible and well-regarded source of certification.

The competition focuses on very early-stage ventures. To be eligible to participate, companies should not yet be incorporated, or only very recently, and cannot have already received funding from professional investors. These requirements imply that Venture Kick is usually involved in

[7] According to the OECD, Switzerland had more than 70 million venture capital dollars invested per one million inhabitant in 2019.

[8] This ranking is based on CrunchBase's "CB Rank (Company)" consulted in April 2019.

the startups' development before any other external investors or entrepreneurship programs. Additionally, participating startups should have links with a Swiss academic institution or research center and plan to establish their company and create jobs in Switzerland. Conditions attached to the grant money have not changed significantly during the years of our study, from 2007 to 2017.[9] The first $10,000 and $20,000 are given as a grant with a moral obligation to repay in case the startup becomes successful. The final $100,000 was initially also given as a grant but this changed in 2013, when Venture Kick started taking a 1 to 3 percent share in the winning companies.

Venture Kick is structured as a three-stage, nine-month, private competition open several times a year to a new group of eight startups. Not all startups that apply are accepted; around 20 percent of applicants do not participate for a mixture of quality and logistical reasons. The competition's judges are drawn from a pool of more than 150 private and institutional investors, startup experts and industry representatives. 27 percent of the judges list Internet & Mobile and Software as their main area of expertise while the rest are experts in more science-based technologies (e.g., biotech and medtech).[10] These judges register for the competitions based on their availability without knowing the participants. In the first stage (stage 1), startups pitch their business idea in front of a jury composed of an average of nine judges. Each jury member grades the startups on eleven different criteria and votes for the startups they deem worthy of advancing to the next stage.[11] These four winning startups receive $10'000 and Venture Kick advertises the startups' victory on its website and social media. Winning startups also gain the right to display the competition's logo on their own website—a logo that is identical whether the startup wins only stage 1 or all three stages.[12] Finally, they secure access to the first two, out of three, kicker camps. These camps are two-day events where entrepreneurs receive mentoring and are put in contact with accomplished entrepreneurs and industry experts.

In the second stage (stage 2), three months later, the startups have to present their business case in a similar setting to a different set of judges and two startups are selected for the third and final stage (stage 3). These two startups receive $20'000 as well as the same written publicity as in stage 1. They also gain access to the last kicker camp. Six months later in stage 3, the last two startups compete to win the competition and the $100'000 prize attached to it. Once more, Venture Kick communicates about the ultimate winner's success. Figure A.1 in Appendix A illustrates the whole process. In addition, after each stage Venture Kick provides successful and unsuccessful participants with private feedback on their performance. This feedback includes a video of their presentation and subsequent Q&A, their grades and a report

[9] In 2019, Venture Kick increased the money given in the second stage from $20,000 to $40,000, increasing the total amount received by the final winner from $130,000 to $150,000. The conditions attached to the funding have also changed. The initial $10,000 still given as a grant but the remaining $140,000 now a convertible loan with an interest rate of 5 percent. As these changes occurred outside of the time frame of our study, we do not look into their implications.

[10] Unfortunately, we cannot link the expertise of the judges to their choice of winning startups as jury participation is anonymised.

[11] Table A.1 in Appendix A shows the exact grading categories given during the three stages.

[12] Appendix A provides concrete examples of this publicity.

that contains all the jury's comments and open questions about their project's viability.

1.4.2 Data provided by Venture Kick

Venture Kick has provided us with anonymized data on its competition and on 987 startups that have participated in a total of 305 rounds of competition over the years 2008 to 2017.[13] The data include information on the startups (e.g., location, link with university, industry, gender of founder) and their performance during and after the competition.[14] These data form one of the strengths of our paper as all the startups have gone through exactly the same process, with the same set of rules, and were motivated by the same financial and non-financial incentives. Also, the relatively low number of participants per competition (a maximum of eight) means that the judges have the time (around 30 minutes) to analyse each startup adequately. As this process is repeated up to three times, each startup that competes in stage 3 will have spent 90 minutes with Venture Kick's judges. This setup contrasts with previous studies. Howell (2020)'s larger sample includes 87 different competitions that each had between six to 275 participants and one to 178 judges. In McKenzie (2017), there is only one paper-based competition during which the 6000 applications are assessed by an algorithm and an evaluator.[15] We exploit the consistency and quality of the grading data in our identification strategy, as explained in Section 1.5.

Venture Kick has recorded every grade and vote given by the jury members in all three stages since 2008 and has shared them with us. In stage 1, each judge grades the startups on ten elements summarized by a 'People' and a 'Project' grade and an eleventh grade denominated 'Gut Feeling.' In stages 2 and 3, the 'Gut Feeling' grade remains but the other ten grades are replaced by ten new grades summarized by a 'People & Project' grade and an 'Achievements & Development' grade. Grades are given on an absolute scale from 0 (worst) to 6 (best). This absolute scale allows us to identify weaker cohorts, providing some useful variation in the quality of both losers and winners. Because the variance in grades given to different startups decreases through the stages as the quality of the participants becomes consistently higher, we standardize each grade at the stage level. For our analysis, we decided to aggregate our eleven grade variables into one single measure by using a principal component analysis (PCA).[16] Table A.2 in Appendix A shows the eigenvalues of the PCA on all the grades given in stage 1. Using the Kaiser Criterion, dropping all components with eigenvalues below 1, we retain only the first principal component that accounts for 76 percent of the variance in grades. We interpret this first component as being a measure of startup quality. Moreover, Table A.3 shows

[13] 123 of these competitions are stage 1 rounds, 123 are stage 2 and 59 are stage 3.

[14] One thing that we do not know is the startups' founding date, and, therefore, cannot include the startups' age in our estimations. However, the participation requirements (i.e., no professional investors already involved and no incorporation) ensure that the startups are at the same stage of development when starting the competition.

[15] The first selection of 6000 applications is made by an algorithm after a rapid assessment (ten minutes approximately). In the second stage the remaining 4510 business plan applications are marked by an evaluator drawn from a team of 20 individuals, with a typical plan taking 30-45 minutes to mark.

[16] We use this approach because we have noticed that the grade variable it produces for each stage is a better predictor of future startup success than any other aggregation method (e.g., average of the grades).

that the first component loads similarly onto all the grades, which is what we would expect as startup quality should be reflected in all the grades. Henceforth, the grade variable we use in the analysis is the measure produced by the PCA, unless specified otherwise. On the startups themselves, we have data on their characteristics including, industry, date of creation, gender of the main founder or the university it is affiliated with).

1.4.3 Additional data collected

Testing the model's predictions requires data on both the decisions of the entrepreneurs and the investors. To capture the entrepreneur's decision to pursue their activity, we use data on startups' survival status in the years after the competition. To capture the investor's decision to invest in a given startup, we use data on whether and how much startups are able to secure external funding in the years after the competition. The data collection process was performed under our close supervision by Venture Kick to preserve the anonymity of the startups' information.

To construct the annual survival variable, we manually verified both whether and when each of the 759 startups that competed between 2008 and 2015 stopped its activity during the period up to the 1st of January 2018.[17] To do so, we relied on the following sources: Cantonal registries of commerce; the startup's website status and activity—if the website was down we used the Internet Archive's Wayback Machine[18] to find out when it stopped being active—; LinkedIn's staff profile and startup pages; Facebook and Twitter activity; CrunchBase's databases and other resources like newspaper articles and Venture Kick's own resources. We followed the same process to obtain data on startup exit (IPO & Acquisitions).[19] To build our funding variables, we manually collected funding information on CrunchBase for the 970 startups that competed between 2008 and 2017 for the period up until the 1st of January 2020.[20] These data provide the number of funding rounds, their dates and the amount of funding received for all participating startups. We also used patent data from the PATSTAT database to measure the startups' innovation output and matched 693 patents to 104 startups in our sample.

Note that having access to annual outcome data is quite rare in the literature. These data will allow us to differentiate between the short-term and long-term effects of the certification and cash prize.

[17]The survival status of the startups that competed in 2016 and 2017 was deemed "too young to tell" because most of them had less than 2 years of existence at the time of the data gathering process (i.e., spring to summer 2018).

[18]The Internet Archive is a San Francisco based non-profit digital library and its WayBack Machine archives the entire web: WayBack Machine.

[19]Unfortunately, as we show in Table 1.1, exit is a very rare event, happening to only 3 percent of the startups in our sample. This lack of variability prevents us from using exit data in our regression analyses.

[20]These data were collected in early 2020 and therefore contain two more years than the survival variable.

1.4.4 Descriptive statistics

Table 1.1 shows descriptive statistics concerning the startups in our sample, depending on how far they progressed into the competition. This table yields interesting facts about participants and winners. The first finding is that software and internet & mobile startups do not perform well in the competition. Although they represent 43 percent of the sample, only 17 percent of the winners of stage 3 come from these two sectors. At the other end of the spectrum, medtech, biotech and electronics & mechanics startups are over-represented among the winners of stage 3 (58 percent). Similarly, two universities dominate the competition. The two federal institutes of technology in Lausanne (EPFL) and Zurich (ETHZ) represent 44 percent of the participants but 72 percent of the winners. Women entrepreneurs are under-represented, with 89 percent of the startups having declared a man as the main founder. However, they exhibit roughly the same success rate as men.

As could be expected, startups that go further in the competition tend to have a higher probability of survival, 27 percent of the startups that lost in stage 1 are alive five years after the competition versus 93 percent for startups that won the final stage (last column). They also tend to secure more financing, have more employees and larger turnover, and obtain more patents in the years following the competition. Overall, the data tell us that winning the competition correlates with better performances after the competition. We now turn to our empirical strategy to explore the causal impact(s) of venture competitions. Additionally, Table 1.1 shows that startups participating in Venture Kick are very similar to their counterparts in other developed countries. Indeed, in the United States, Howell (2020) finds that 3 percent of the startups that participated in the studied venture competitions have exited, which is the exact same number as in our sample. Moreover, 24 percent of these American startups secured funding after their competition compared to 26 percent after Venture Kick. The similarities between the startups in our sample and in similar studies bolster the external validity of our findings.

1.5 Empirical strategy

Our main empirical challenge lies in estimating the causal effect of winning each of Venture Kick's stages. We explain below how we leverage the effect of winning stages 1 and 3 to capture the effect of certification and cash prize.

We test for the presence of the certification effect using stage 1 because it is mainly this first stage that provides information about a startup's quality to entrepreneurs and investors. Indeed, as Venture Kick is intended for very early-stage startups that professional investors have not yet scrutinized, stage 1 plays a crucial certifying role. Participating in stage 1 provides startups with feedback that is very often the first expert assessment of their potential, encouraging winners to pursue their ventures while discouraging losers. In addition, winning stage 1 usually provides startups with the first credible and informative signal of their quality

Table 1.1: Descriptive statistics: Characteristics and Outcome

	Full sample	Lost in Stage 1	Lost in Stage 2	Lost in Stage 3	Winners
	Mean	Mean	Mean	Mean	Mean
Categories:					
Software	0.24	0.31	0.26	0.16	0.08
Internet & Mobile	0.19	0.20	0.20	0.18	0.09
Medtech	0.12	0.07	0.13	0.17	0.23
Electronics & Mechanics	0.10	0.05	0.11	0.14	0.19
Biotech	0.09	0.04	0.11	0.15	0.16
Cleantech	0.07	0.04	0.09	0.10	0.09
Materials & Chemicals	0.03	0.03	0.04	0.03	0.02
Micro-Nano Technology	0.03	0.01	0.02	0.06	0.07
Others	0.05	0.09	0.04	0.02	0.02
Universities:					
EPFL	0.21	0.13	0.22	0.31	0.33
ETH	0.23	0.17	0.25	0.26	0.39
Miscellaneous:					
Main founder is a man	0.89	0.86	0.94	0.88	0.88
Average grade in Stage 1	3.85	3.32	4.22	4.34	4.49
Average grade in Stage 2	3.97	.	3.56	4.25	4.40
Average grade in Stage 3	4.27	.	.	3.89	4.62
Outcomes:					
Alive after 2 years	0.61	0.36	0.70	0.84	1.00
Alive after 5 years	0.52	0.27	0.59	0.73	0.93
Received Funding after 2 years	0.19	0.06	0.19	0.30	0.46
Received Funding after 5 years	0.22	0.07	0.19	0.35	0.62
No. of Funding rounds after 2 years	0.40	0.20	0.25	0.53	0.72
No. of Funding rounds after 5 years	0.81	0.48	0.51	0.93	1.40
No. of patents after 2 years	0.22	0.04	0.15	0.28	0.51
No. of patents after 5 years	1.74	0.15	1.49	2.42	3.05
No. of Employees after 2 years	7.06	.	5.24	7.52	9.64
No. of Employees after 5 years	14.70	.	11.39	13.71	20.63
Turnover after 5 years, kCHF	810.85	.	351.12	656.98	1623.60
Acquired/IPO'ed	0.03	0.01	0.04	0.09	0.05
Observations	983	433	267	129	121

Note: This table presents descriptive statistics for the full sample and conditional on their results in the competition. Number of Rounds, Patents, Employees, and Turnover are reported for the sub-sample of startups that are still alive. The total number of observations for the first eleven outcome variables are respectively 713, 428, 945, 606, 432, 224, 468, 234, 331, 176 and 171.

for outside investors. This signal takes the form of an award and some online publicity.[21] It also comes with increased online visibility—97 percent of the startups that won stage 1 have a profile on CrunchBase compared to only 23 percent of the losers. Furthermore, stage 1's $10,000 cash prize is not consequential.[22] To put this number in perspective, the median monthly salary in Switzerland stood at around $6,500 in 2019.

Winning stage 2 is mainly about reinforcing the certification and learning effect of Venture Kick. Upon winning stage 2, startups receive $20,000 and some more publicity. The money remains limited, and the publicity is somewhat redundant, reinforcing the certification effect.[23]

We use stage 3 to test for the presence of the cash prize effect: winning stage 3 yields $100,000—more than three times the combined prizes for winning stages 1 and 2—raising the total funding received to $130,000. While startups might also benefit from being declared the ultimate winner of the competition, we argue that stage 3's certification is of much less significance than stage 1's. For one, it is unlikely that, after having received positive comments in stages 1 and 2, the feedback received during Venture Kick's final stage will have much bearing on entrepreneurs' decision to pursue their venture. On top of that, as the two finalists will both be of high quality, winning stage 3 is not as informative a signal to investors as winning stage 1 where there is much greater heterogeneity in participants' quality.

It is important to acknowledge that Venture Kick also offers other value-added activities during all three stages. Indeed, the judges' feedback and the competition in itself provide participants with opportunities to increase their network and learn how to improve their ideas or products. However, these aspects of the competition are not restricted to one particular stage, making their estimation difficult. Furthermore, they are offered to both the winners and losers of each stage. Consequently, because the networking and learning opportunities lift all boats, the presence of these unobserved value-added activities does not threaten our identification strategy.

To estimate the treatment effect of winning each stage, we must deal with the fact that higher quality startups are more likely to win. The empirical challenge lies in disentangling the treatment effect from the 'startup quality' effect. We address this challenge by using detailed data on grades. As we document below, judges' grades can be used to capture the startups' intrinsic quality. Controlling for startup quality, we can then isolate the treatment effect of winning a stage of the competition. Additionally, we implement alternative identification strategies that do not rely as much on judges' grading. Notably, we use a sharp regression discontinuity design that exploits the discontinuity in treatment between the startups ranked fifth, the best losers, and the startups ranked fourth, the worst winners.

[21] To view some endorsement and details of Venture Kick's publicity see Appendix A.

[22] One participant provided this quote on the matter: *"If you look back the $10,000 of the first Venture Kick award seems not much."*

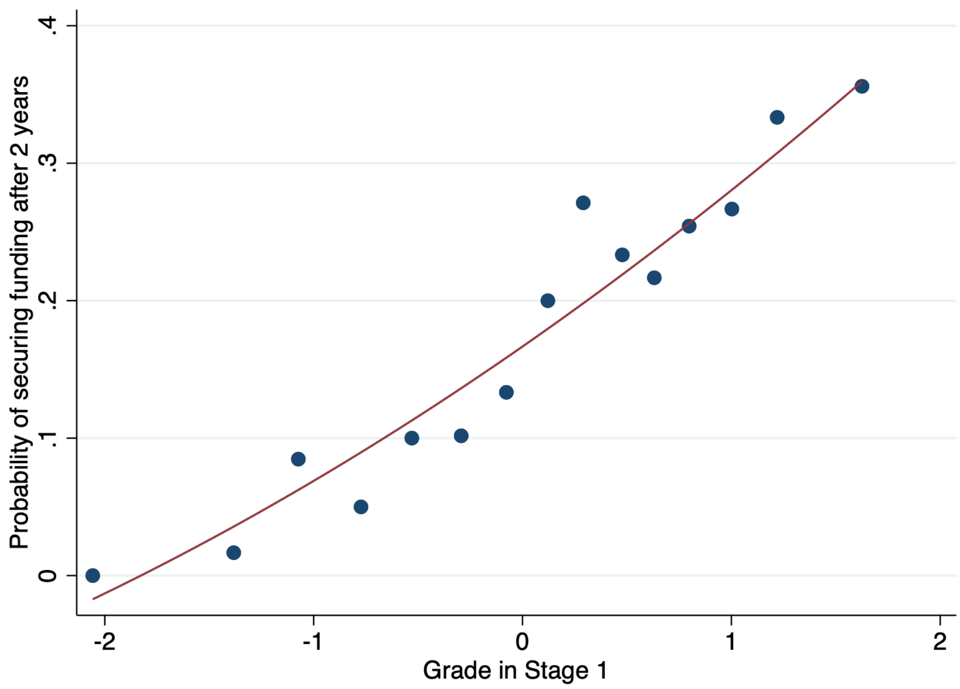
[23] While we do not show any estimate of stage 2's treatment effect in this paper, we find that Venture Kick's second stage does not have any impact on startups' future business success. One potential reason behind this result could be that stage 2 simply lacks any distinctive reward that might significantly affect a startup's performance.

The next section provides evidence that grades correlate with startup quality.

1.5.1 Grades reflect startups' quality

The grades given by Venture Kick's judges are particularly useful because each startup in our sample has been assessed by the same type of jury with the same motivations and guidelines. Moreover, we argue that these grades are an informative proxy for startup quality because they strongly predict a startup's success both in later stages and after the competition. Looking back at Table 1.1, we see that judges can, on average, already predict in stage 1 who the eventual winners of the competition will be. The average grade they give during the first stage to the startups that will eventually lose in stage 2 (4.22) is lower than for losers of stage 3 (4.34) and final winners (4.49). This cannot be due to favoritism or judges sticking with their first choice because the jury changes at each stage. More importantly, judges' ability to predict success also applies to post-competition performances. Figure 1.1 shows that there is a clear positive relationship between judges' grades in stage 1 and a startup's probability of raising subsequent external funding in the two years after the competition.[24]

Figure 1.1: Judges are able to accurately discern startups' true quality



Note: This figure shows the strong positive relationship between judges' grades in stage 1 and the startup-specific probability of securing funding in the two years after the competition.

More formally, Table 1.2 displays the linear relationship between the grades given in stages 1,

[24] Note that because grades are never made public, investors cannot condition on grades when deciding on funding.

2 and 3 and survival two to eight years after the competition.[25] We see that these grades are a reliable predictor of future survival rate, especially in the short run. A one-standard-deviation increase in grade in stage 1 is associated with an increase in survival rate three years after the competition of 23 percentage points. In the longer run, as the sample size decreases and startups evolve, this relationship understandably fades. Stage 3's lower coefficients in the first three years of the competition are explained by the minimal variation in the dependent variables. Indeed, 100 percent of the winners of stage 3 are alive after two years, while 96 percent of them are still alive after three and four years. Judges show the same capability to predict success if we use funding variables as outcomes or control for winning stages of Venture Kick. In a nutshell, we argue that, because judges' grades are an informative predictor of future success, we can use them as a proxy to control for the quality of each startup.

Table 1.2: Relationship between judges' grades and future success

	(1) Survival 2	(2) Survival 3	(3) Survival 4	(4) Survival 5	(5) Survival 6	(6) Survival 7	(7) Survival 8
Panel A: Stage 1							
Standardized grade	0.198***	0.222***	0.202***	0.185***	0.144***	0.117**	0.0439
	(0.0271)	(0.0285)	(0.0307)	(0.0325)	(0.0421)	(0.0533)	(0.0662)
Panel B: Stage 2							
Standardized grade	0.128***	0.0943**	0.0757*	0.0688	0.0302	0.0540	0.0711
	(0.0340)	(0.0401)	(0.0431)	(0.0514)	(0.0485)	(0.0596)	(0.0775)
Panel C: Stage 3							
Standardized grade	0.0789**	0.0814**	0.127***	0.152**	0.134*	0.135	0.109
	(0.0339)	(0.0349)	(0.0419)	(0.0631)	(0.0773)	(0.116)	(0.132)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Competition FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
University FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations stage 1	657	555	457	387	297	211	119
Observations stage 2	354	305	256	214	172	123	77
Observations stage 3	179	157	135	114	95	70	48

Note: This table presents results of linear regressions of survival status on grades received during the stages 1 through 3. The dependent variable is the survival status two years after the competition, *Survival 2*, in Column (1) all the way to eight years after the competition, *Survival 8*, in Column (7). The grade variable is standardized at the stage level. The level of observation is a startup in a competition round. Errors are clustered by competition round. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

However, the ability of Venture Kick's judges to objectively assess each startup's quality varies depending on the startup's technology, in line with previous findings in the literature (Scott et al., 2020). Our data show that Venture Kick's experts are better at discerning between high- and low-quality early-stage startups in the hard science sectors (e.g., medtech, biotech, materials & chemicals) than in the ICT sectors. The top panel of Figure 1.2 shows the coefficients of the regressions in Table 1.2 using interactions terms to differentiate between

[25] We do not control for winning the stage at this point because we want to focus on the grades. Moreover, the relationship between grades and future performance remains the same whether we control for it or not.

two technology subgroups (ICT startups vs. all other remaining startups, which are mostly science-based startups).[26] In the bottom panel, the outcome variable captures whether or not a startup managed to secure funding after the competition. Overall, Venture Kick's jury excels at evaluating the quality of science startups, and its grading remains informative in the long run. However, the grading received by an ICT startup is a weaker predictor of future business success. Similarly, the standard deviation in the project grades given to ICT startups during stage 1 by Venture Kick's judges is almost 4 percent higher than the standard deviation in the project grades given to science-based startups.[27] The higher variation in grades could reflect a greater difficulty in evaluating ICT startups. This heterogeneity in the judges' grades predictive power and dispersion is important. Indeed, an institution's certification of quality relies on its ability to correctly assess said quality. Because it is harder for experts to objectively evaluate ICT startups, Venture Kick's certification of their quality is less trustworthy than for science startups. Therefore as we hypothesized in Section 1.3.1, winning stage 1 should be more valuable for science ventures than for software-focused ones.

1.5.2 Empirical specification

Having established that grades strongly correlate with a startup's future success, we use them to estimate the causal effect of the certification effect (winning stage 1) and the cash prize effect (winning stage 3). First we estimate these effects over the whole sample by implementing the following regression:

$$Y_{i,s} = \alpha_{i,s} + \delta_1 WinsStage_{i,s} + \beta_1 StartupQuality_{i,s} + \gamma_{year_s/comp/ind/uni} + \epsilon_{i,s} \qquad (1.1)$$

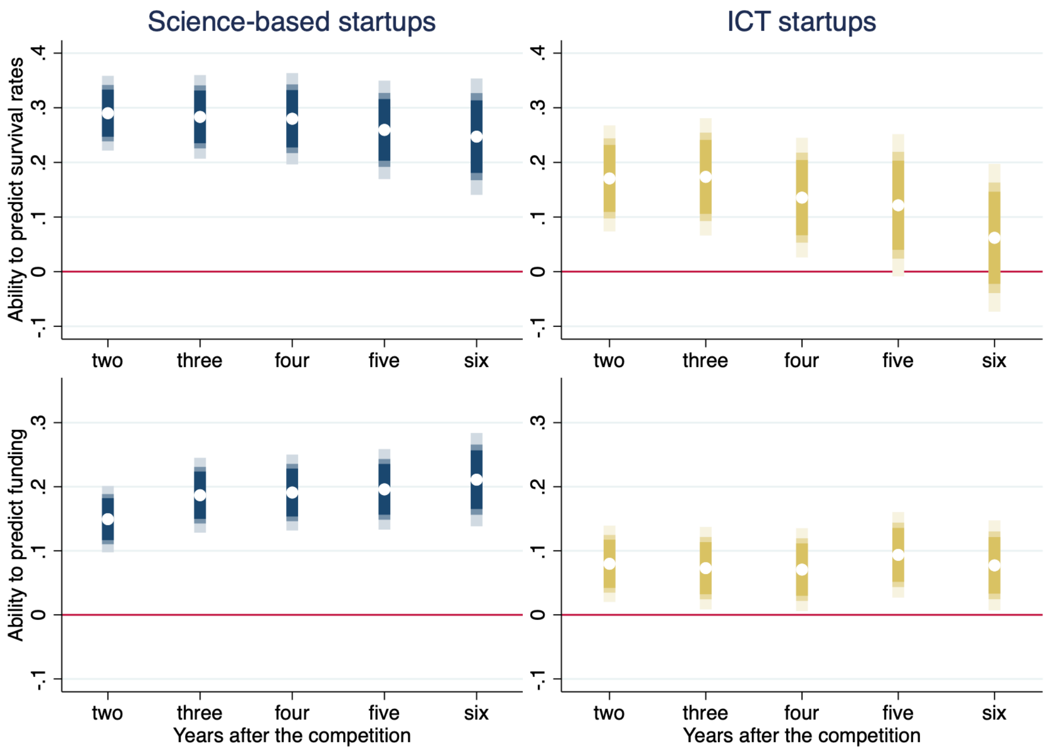
where i represents the startup and s the stage. The dependent variable $Y_{i,s}$ is either a survival or a funding outcome. The *Survival 2* to *Survival 8* variables take a value of 1 if a startup is alive two to eight after the competition and a 0 otherwise. *Funding 2* to *Funding 8* are also binary variables that take the value of 1 if a startup secures non-Venture Kick funding two to eight years after the competition. In some specifications, we also look at *Funding*, which does not impose any time restriction on receiving funding.[28] Finally, *Rounds 2* to *Rounds 8* are the number of funding rounds that a startup manages to secure two to eight years after the competition and take values from 0 to 10. Fixed effects for year (γ_{year_s}), competition (γ_{comp}), industry (γ_{ind}) and university (γ_{uni}) are also added. The coefficient δ_1 measures the treatment effect of winning each stage whereas $StartupQuality_{i,s}$ is captured by the

[26]The startups classified as science-based fall in the following categories: medtech, electronics & mechanics, biotech, cleantech, materials & chemicals, micro-nano technology and others. ICT startups are those in the internet & mobile and software categories.

[27]We focus on the project grades as these directly capture the quality of the project and its technology. Judges do not, and should not, differ in their ability to judge people and teams in the ICT and science-based sectors.

[28]We use both the time-restricted and unrestricted *Funding* variable because both have their own appeal. Focusing only on the first two years, for example, prevents our funding outcome from being too influenced by the startups' survival rates, as even the weakest startups will be alive and able to search for financing for a good part of these two years. At the same time, focusing on all the available data puts startups whose treatment effect might take longer to materialize on a more equal footing.

Figure 1.2: Relationship between a one-std.-deviation increase in grades and future outcomes



Note: This figure's top panels displays the coefficients of Table 1.2 (i.e., how well the judges' grades predict survival two years after the competition) using interactions terms to differentiate between technology subgroups. The bottom panels show the relationship between grades and whether startups have secured financing after the competition. Errors are clustered by competition round. The confidence intervals are at the 10%, 5% and 1% levels, respectively.

startup's grades. Given the data structure and the fact that one firm's win is another firm's loss, errors are clustered by competition round. We use OLS as our main specification.[29] As an extension, we add control variables and implement errors-in-variables regression models to account for the fact that grades might be an imperfect proxy of startup quality.

Equation (1.1) allows us to estimate the average treatment effect of venture competition but it is silent on industry-specific effects. We therefore augment equation (1.1) by including interaction terms to test our predictions that venture competitions' rewards have a different effect depending on the startups' characteristics.

We divide the startups in our sample into two technology groups: science-based and ICT. This separation allows us to identify the impact of the competition on $high\_info$ and low_info startups. As discussed in Section 1.5.1, judges are better able to assess the quality of science-

[29]We have also estimated logit models and obtained similar findings. We report OLS regression results because they are easier to interpret. Furthermore, OLS does as well as logit, if not better, in estimating causal effects of treatment (Angrist, 2001).

based startups ($high\_info$) than ICT startups (low_info). Moreover, startups in these two technology types typically have different levels of running costs. Science-based startups can require significant capital investments while ICT startups have lower costs, in part because they have recently benefited from technological shocks that have significantly reduced the cost of running a business in these sectors (e.g., the advent of cloud computing as argued by Ewens et al., 2018).

To estimate the differentiated treatment effect of each stage on ICT startups and science-based startups, we use equation (1.2) below:

$$\begin{aligned} Y_{i,s}^{Hetero} = {} & \alpha_{i,s} + \delta_1 WinsStage_{i,s} + \delta_2 WinsStage_{i,s} * ICT_i + \beta_1 StartupQuality_{i,s} \\ & + \beta_2 ICT_i + \beta_3 StartupQuality_{i,s} * ICT_i + \gamma_{year_s/comp/uni} + \epsilon_{i,s} \end{aligned} \tag{1.2}$$

Compared to equation (1.1), we now allow for a different effect of winning each stage between sectors. The coefficient δ_1 is now the effect of winning a given stage for science-based startups and δ_2 captures the difference in effect between ICT and science startups. Consequently, the effect of winning for ICT startups is the sum of δ_1 and δ_2. We also allow for a different relationship between judges' grades and future outcomes between sectors (β_3)

1.6 Results

1.6.1 Venture Kick's certification effect

Certification impacts survival and funding

Using equations (1.1) and (1.2) we estimate the treatment effect of winning stage 1, our test for the certification effect, and provide the results in Table 1.3. Focusing first on the competition's average effect in Panel A, column (1) shows that without any controls, winning stage 1 is associated with a 44-percentage-point increase in the probability of survival two years after the competition. This result reflects both the effect of the treatment and startup quality on startup survivability. Column (2) introduces year, competition, industry and university fixed effects. Column (3) adds the grade variable to capture the quality of the startups and is our main specification. We find that winning stage 1 increases survival after two years by 22 percentage points. This is a rather large effect, given that the average survival rate of a Swiss startup two to three years after its launch is between 60 and 70 percent (see Appendix A). Columns (4) and (5) respectively present the effect of winning on the probability of receiving external funding after the competition and on the number of funding rounds concluded in the two years following the competition. Here, we look both at funding outcomes with and without the two-year restriction to show that results do not depend on this parameter. As with survival rates, winning stage 1 has a significant positive impact. It increases the probability of securing financing after the competition by 14 percentage points and increases the number of funding rounds within two years by 0.25. To put these numbers into perspective, if the average startup in our sample wins stage 1, its probability of securing financing after the competition rises from 18 percentage point to 32 percentage point, a nearly 80 percent increase. As with survival, the treatment effect of winning stage 1 on subsequent financing is quite large.

The certification effect varies across technologies

Results in Panel A tell us that, on average, certification improves startups' business success. However, these average effects conceal the fact that Venture Kick might have a heterogeneous effect on participating startups. Panel B addresses that limitation. It presents the differential effect of Venture Kick on ICT and science-based startups. Focusing on columns (3) to (5), it appears that winning stage 1 has a positive and significant impact on science-based startups business success. It increases their survival rate two years after the competition by 41 percentage points. Winning stage 1 also causes the expected number of funding rounds secured within two years to rise by 0.4 and the probability of securing funding by 22 percentage points. However, as we predicted, the certification effect is significantly reduced for ICT startups. The difference in effect is such that ICT startups do no see any significant benefit from winning stage 1 (see Figure 1.3). These results confirm that Venture Kick's certification effect varies across sectors, in line with Proposition 1. The certification has significantly less impact on the outcome of ICT startups than it has on the outcome of science-based startups, whose quality can be more objectively assessed.

The certification effect is long-lasting

We now turn to studying how persistent the certification effect is over time. We predict that because the information provided by the certification does not disappear, the certification effect should be long-lasting (Proposition 1.a). To test this prediction, we simply estimate the treatment effect of winning stage 1 on science-based startups and ICT startups two to six years after the competition, using equation (1.2). The estimates are displayed in Figure 1.3, which complements the results from Table 1.3. They show that the certification has a long-lasting effect on science-based startups' survivability and funding prospects. Winning stage 1 increases survival rates between 33 and 41 percentage points in the six years after the competition. For ICT startups, expanding the time horizon does not change the conclusion: Venture Kick's certification does not affect survival and funding.

Losing entrepreneurs are most impacted by the certification

Finally, we delve deeper into the exact nature of the certification effect. Because Venture Kick's certification impacts both the winners and the losers of each stage, we cannot directly know whether the positive treatment effect of winning stage 1 comes from a positive effect on the winners or a negative effect on the losers. To study this question, we compare the predicted survival rate of the average startup in our sample with the average survival rates of young Swiss

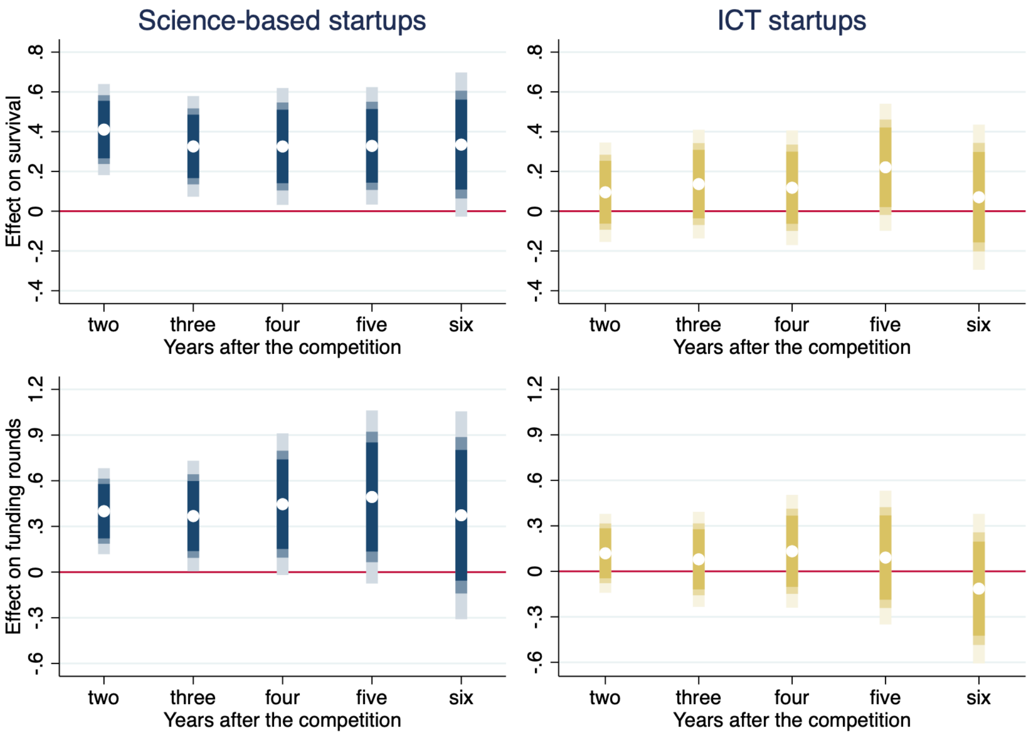
Table 1.3: Main specification: Effect of winning stage 1, the *certification effect*

	Survival after the competition			Funding after comp.	
	(1) Survival 2	(2) Survival 2	(3) Survival 2	(4) Funding	(5) Rounds 2
Panel A: Average effect					
Wins Stage 1	0.437***	0.373***	0.220***	0.140***	0.249***
	(0.0351)	(0.0491)	(0.0683)	(0.0454)	(0.0890)
Standardized grade			0.119***	0.0919***	0.134***
			(0.0353)	(0.0252)	(0.0463)
Panel B: Science vs ICT					
Wins Stage 1	0.572***	0.593***	0.410***	0.217***	0.400***
	(0.0450)	(0.0570)	(0.0870)	(0.0539)	(0.108)
Wins Stage 1 x ICT	-0.304***	-0.333***	-0.315**	-0.153**	-0.281**
	(0.0677)	(0.0906)	(0.123)	(0.0753)	(0.129)
ICT	0.112**	0.127**	0.142*	0.0583	0.105
	(0.0517)	(0.0598)	(0.0792)	(0.0478)	(0.0807)
Standardized grade S1			0.137***	0.136***	0.156***
			(0.0404)	(0.0294)	(0.0550)
Grade x ICT			-0.0155	-0.0805**	-0.0678
			(0.0571)	(0.0338)	(0.0604)
Year FE	No	Yes	Yes	Yes	Yes
Competition FE	No	Yes	Yes	Yes	Yes
University FE	No	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes	Yes
Observations	671	671	671	893	893
R^2	0.196	0.397	0.413	0.307	0.321

Note: This table presents estimates of the effect of winning stage 1 and grading on whether the startups survived or secured financing a given number of years after the competition.Panel A displays the average effect of winning stage 1 using equation (1.1), and Panel B displays the sector-specific effects using equation (1.2). Columns (1) through (3) use the survival status two years after the competition as the dependent variable, with varying levels of fixed effects. Column (4)'s dependent variable is whether the startup secured financing after the competition while Column (5)'s is the number of funding rounds secured within the two years of the competition. The level of observation is a startup in a competition round. Errors are clustered by competition round. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

and American ventures. Appendix A presents the full analysis. Comparing participating and non-participating startups, it seems that winning does not significantly impact survival rates. Indeed, the predicted survival probability of the average Venture Kick startup if it wins stage 1 is very close to the average survival rate of all Swiss startups. However, the predicted survival rate of this same average startup if it loses stage 1 is much lower than the national average. This finding is consistent with the idea that losers, discouraged by negative feedback, cut their

Figure 1.3: Effect of winning stage 1 on survival and funding over the years for two broad technology categories



Note: This figure presents the effect of winning stage 1 on *Survival* (top panel) and *Rounds* (bottom panel) for science-based startups and ICT startups, two to six years after the competition using Equation 2. Errors are clustered by competition round. The confidence intervals are at the 10%, 5% and 1% levels, respectively.

losses much quicker than they would have, had they not participated in the competition.

1.6.2 Venture Kick's cash prize effect

The cash prize increases ICT startups' survival rates...

We now study Venture Kick's cash prize effect by estimating the treatment effect of winning the third stage. Table 1.4 displays the average effect of winning stage 3 using equation (1.1) in Panel A and the technology-specific effect using equation (1.2) in Panel B. Interestingly, and in contrast to stage 1, winning stage 3 does not seem to have any effect on startups' future business success on average. However, once again this effect hides sector-specific differences. As column (3) in Panel B shows, winning stage 3 has no impact on the survival of the science-based startups. This result is not surprising as the \$100,000 prize alone covers a science startup's expenses for a limited period of time. However, as posited in Proposition 2, startups with lower running costs can significantly extend their runway by winning Venture Kick's cash prize. Winning stage 3 and the related \$100,000 prize increases the survival rate of ICT ventures two years after the competition by 23 percentage points.[30] Finally, columns (4) and (5) tell us that the cash prize has no effect on both science-based and ICT startups' probability of securing funding after the competition. The fact that investors are not affected by the cash prize is unsurprising, given that they cannot claim the prize for themselves and that the extra money does not fundamentally change a startup's quality.

... but this effect is temporary

Our second prediction about the cash prize effect is that it is short-lived (Proposition 2.a). Indeed, if the cash prize is helpful because it extends startups' runway, then the cash prize should not have a significant impact once the prize money is exhausted. To test this prediction, we estimate the cash prize effect over the years. Figure 1.4 displays the results, focusing on survival (the cash prize never has any impact on funding). It confirms that the cash prize has no impact on science-based startups but, more importantly, provides an insightful visualization of the timing of the cash prize effect on ICT startups. The data suggest that winning the whole competition and the \$100,000 cash prize only causes an increase in survival rates during the first two to three years that follow the event. After that, a reasonable period for an ICT startup to consume the \$100'000 prize, there is no sign that having received a cash prize has any impact on startups' survival.

1.7 Robustness tests

We now address potential concerns about our main identification strategy. We first confirm that our results are robust to including control variables and to using different versions of the grade variables. We then test the sensitivity of the results to variations in the grades' reliability by estimating errors-in-variables regression models. Finally, we implement two alternative identification strategies that do not rely so much on the grades. Most notably, we use a regression discontinuity design (RDD) that exploits the fact that the startups ranked fourth during stage 1 receive treatment while those ranked fifth do not. In addition, we implement a propensity score matching (PSM) analysis.

1.7.1 Introducing additional control variables

This section shows that our findings are robust to the inclusion of additional control variables. First, we control for motivation, which, like quality, could impact a startup's performances both during and after Venture Kick. We proxy motivation by looking at whether startups participated in the optional kick camps offered by Venture Kick. Second, Clingingsmith

[30] This ICT-specific effect is not enough to drive a positive average effect because, while ICT startups represent 43 percent of our sample, they only represent 17 percent of stage 3 winners. This result therefore cautions against using average treatment effects when analyzing the impact of venture competitions such as Venture Kick. Indeed, in our case, they hide the fact that cash prizes are actually useful for ICT startups, a growing share of the startup universe.

Table 1.4: Main specification: Effect of winning stage 3, the *cash prize effect*

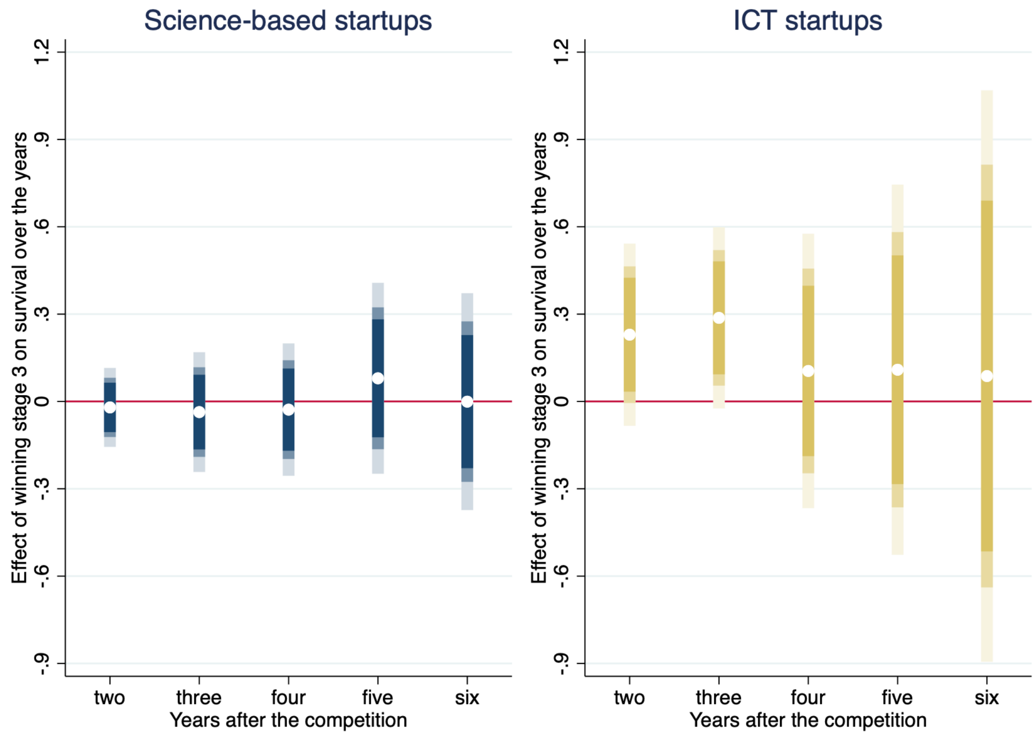
	Survival after the competition			Funding after comp.	
	(1) Survival 2	(2) Survival 2	(3) Survival 2	(4) Funding	(5) Rounds 2
Panel A: Average effect					
Wins Stage 3	0.105***	0.117**	0.0477	0.171	0.262
	(0.0311)	(0.0476)	(0.0479)	(0.139)	(0.311)
Standardized grade			0.0578	0.0733	0.137
			(0.0413)	(0.0807)	(0.166)
Panel B: Science vs ICT					
Wins Stage 3	0.0357	0.0483	-0.0205	0.142	0.195
	(0.0250)	(0.0486)	(0.0506)	(0.163)	(0.375)
Wins Stage 3 x ICT	0.198**	0.287*	0.250*	0.113	0.0533
	(0.0853)	(0.146)	(0.125)	(0.221)	(0.555)
ICT	-0.198**	-0.236*	-0.223**	-0.0203	0.325
	(0.0853)	(0.120)	(0.101)	(0.132)	(0.321)
Standardized grade S3			0.0562	0.0513	0.203
			(0.0550)	(0.0974)	(0.190)
Grade x ICT			0.0470	0.113	-0.155
			(0.0727)	(0.116)	(0.285)
Year FE	No	Yes	Yes	Yes	Yes
Competition FE	No	Yes	Yes	Yes	Yes
University FE	No	Yes	Yes	Yes	Yes
Industry FE	No	Yes	Yes	Yes	Yes
Observations	178	178	178	216	216
R^2	0.057	0.470	0.485	0.522	0.557

Note: This table presents estimates of the effect of winning stage 3 and grading on whether the startups survived or secured financing a given number of years after the competition. Panel A displays the average effect of winning stage 1 using equation (1.1), and Panel B displays the sector-specific effects using equation (1.2). Columns (1) through (3) use the survival status two years after the competition as the dependent variable, with varying levels of fixed effects. Column (4)'s dependent variable is whether the startup secured financing after the competition while Column (5)'s is the number of funding rounds secured within the two years of the competition. The level of observation is a startup in a competition round. Errors are clustered by competition round. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

and Shane (2017) have shown that founders who pitch their startups first or second in a competition are evaluated more harshly than the others. Hence, we control for the order in which the participants pitched to the jury. Third, we use another grade variable—the average of the eleven grades received—to verify that our results do not depend on the use of a PCA.

Table A.4 in Appendix A shows that including these various controls does not change our result about stage 1's certification effect. The variable *Motivation* is positively associated with

Figure 1.4: Effect of winning stage 3 on survival over the years for two broad technology categories



Note: This figure presents the effect of winning stage 3 on *Survival* for science-based startups and ICT startups, two to six years after the competition using Equation 2. Errors are clustered by competition round. The confidence intervals are at the 10%, 5% and 1% levels, respectively.

future success, but the differential effect of winning stage 1 persists. Column (2) shows that controlling for the order of passage does not change our findings. Column (3) confirms that our results do not depend on which grade we use. We see, however, that the PCA grade in Table 1.3 has a slightly stronger relationship with survival than the average grade in column (3). Finally, including controls does not change the effect of stage 1 on funding outcomes.

We apply the same robustness checks to the effects of stage 3 (see Table A.5 in Appendix A). Our result remains unchanged: the cash prize only has an impact on ICT startups' survivability.

1.7.2 Allowing for grades to be imperfectly observed

As our main identification strategy relies on judges' grading, we want to test the sensitivity of our results to changes in grades' reliability. Indeed, even though the grades capture the best guess of nine experts about the startup quality, these grades remain an imperfect proxy of quality. Accounting for the fact that quality is observed imperfectly using errors-in-variables regression models, our conclusions remain unchanged. Table A.6 in Appendix A shows that assuming that the grade variable has a reliability of 0.7—meaning that the noise represents 30 percent of the total variance—stage 1 continues to be the only stage with a significantly positive average effect.

Going further, the fact that judges are less adept at screening ICT startups means that grades are a relatively less accurate proxy for ICT startups' quality than for science startups' quality. To address this concern, we allow ICT startups' grading to be 20 percent less reliable than science startups'. Comparing columns (2) and (3) of Table A.7 in Appendix A shows that, if anything, assuming more imprecision in ICT grading reinforces our finding. Winning stage 1 has no effect on ICT startups, while stage 3 does not affect science-based ventures. While this approach has its limits,[31] it shows that our results are robust to allowing for some level of imprecision in judges' grading.

1.7.3 Exploiting the sharp discontinuity in treatment (RDD)

Next, we assess the robustness of our results by implementing a regression discontinuity design (RDD). We exploit the sharp discontinuity between the startups ranked fourth and fifth in the first stage, the former going on to the second stage and the latter being eliminated from the competition. We use the startup's rank as the running variable. While having only eight startups per competition raises the issues of the discreteness of the running variable, a setting with eight competitors can still provide useful information, as shown by Howell (2017)'s study on grant competitions with an average of ten applicants.[32] Figure 1.5 confirms that the use of an RDD is warranted, at least in stage 1. We can see on the top-left panel that there is an apparent discontinuity in survival rates and funding probabilities between the startups ranked fourth and fifth in stage 1. The other four panels show that there is no discontinuity around the cutoff in the probability of being an ICT startup, being the fifth startup to pitch, having a man as the principal founder, or being linked to EPFL.[33] There is also no discontinuity at the cutoff for the other pre-competition characteristics. Our design, therefore, satisfies the RDD requirement that there should not be any discontinuity in the other baseline covariates that could cause a discontinuity in the outcome variable of interest. Finally, for an RDD to be valid, treatment should not cause rank. This requirement is easily met because the winners are determined after the ranks are established. Note that because the ranks around the cutoff are equally populated by construction, there is no need to show a density test à la McCrary (2008).

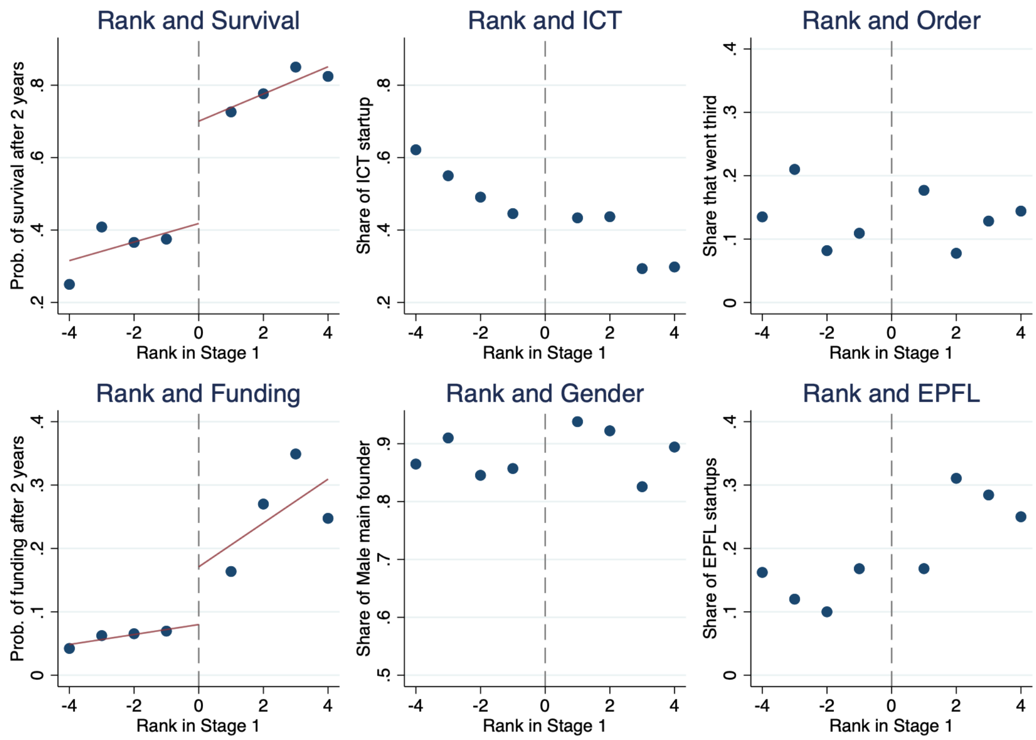
Table 1.5 displays the results of our RDD for stage 1. Columns (1) to (4) are parametric RDDs using the startup's rank in the competition as the running variable, allowing for a different slope coefficient on each side of the cutoff. Columns (5) and (6) are non-parametric RDDs with

[31] We are forced, for example, to drop most of our fixed effects due to the limited number of observations in our sample.

[32] Unfortunately, we cannot implement the RDD in stage 3 due to the reduced number of participants on each side of the cutoff. This is the reason why the RDD is not our preferred identification strategy.

[33] EPFL is one of the two largest research-intensive technical universities in Switzerland.

Figure 1.5: RDD: Discontinuities around the cutoff in stage 1



Note: This figure presents the (dis-)continuity of both our two main outcome variables as well as four selected baseline covariates around the cutoff. The values on display are the probabilities of observing startup-specific outcomes and characteristics at the time of the competition by rank in the first stage. The sample consists of all ventures in the first stage and the result holds for all covariates.

a bandwidth of one rank, meaning that in stage 1 we only focus on the startups ranked fourth (best loser) and fifth (worst winner). We can see that the main results hold: stage 1 and its certification effect only has a positive and significant causal impact on science-based startups' future performances. The size of the certification effect also remains similar. Winning stage 1 increases science-based startups' survival two years after the competition by 45 percentage points and increases the probability of getting funded by 17 percentage points and the number of rounds secured by 0.45. Winning stage 1 still has no effect on ICT startups. Columns (5) and (6) shows the non-parametric results as well as their limits. While column (5) hints at a positive effect for ICT startups, this result is most likely a sign that with only eight possible ranks, startups are not really randomly distributed around the cutoff (fourth and fifth rank). Thus, the non-parametric RDD does not fully solve the endogeneity. Column (6) confirms the results from columns (3) and (4) but the low number of observations is likely causing the interaction term to not be statistically significant.[34] Finally, one general caveat about the use of a RDD in the context of our study is that we cannot implement such design for stage 3 as

[34]This fall in statistical power, illustrates the trade-off between using a larger sample size with larger power and what could be a more precise test at the cutoff that causes a loss in power.

only two startups compete against each other.

Table 1.5: RDD: Effect of winning stage 1

	Sample: All				one rank around cutoff	
	(1) Survival 2	(2) Survival 2	(3) Rounds 2	(4) Funding 2	(5) Survival 2	(6) Rounds 2
Wins Stage 1	0.429***	0.446***	0.445***	0.177**	0.394***	0.218*
	(0.0954)	(0.119)	(0.138)	(0.0692)	(0.101)	(0.122)
Wins Stage 1 x ICT	-0.312***	-0.347***	-0.384***	-0.144**	-0.0918	-0.187
	(0.0722)	(0.101)	(0.130)	(0.0636)	(0.146)	(0.178)
Losers' Rank	0.0336	0.0283	0.0156	0.00867		
	(0.0279)	(0.0329)	(0.0259)	(0.0148)		
Winners' Rank	0.0267	0.0364	0.0374	0.0275		
	(0.0221)	(0.0264)	(0.0334)	(0.0172)		
Year FE	No	Yes	Yes	Yes	No	No
Competition FE	No	Yes	Yes	Yes	No	No
University FE	No	Yes	Yes	Yes	No	No
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	603	603	804	804	172	224
R^2	0.224	0.372	0.320	0.305	0.127	0.019

Note: This table presents results of linear regressions using regression discontinuity designs (RDDs). The running variable is a startup's rank within its competition round. A higher rank is better; in stage 1, -4 is the worst and 4 the best. Columns (1) to (4) display parametric RDDs using the entire sample. Columns (5) and (6) are non-parametric RDDs comparing startups one rank around the cutoff.The level of observation is a startup in a competition round. Errors are clustered by competition round. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

1.7.4 Matching losers and winners (PSM)

The last approach to establish the robustness of the findings involves using a propensity score matching (PSM) estimation. This technique allows us to compare similar startups whose only significant observed difference lies in whether or not they have won a given stage of Venture Kick. We match startups one to one to their nearest neighbor based on the propensity scores. This score is computed using the startups' quality, proxied by the grade variable, their field of technology, whether they have links with EPFL or ETHZ, whether their main founder is a man and the year of the competition. If our matching process is adequate, the matched treatment and control groups should have balanced covariates and propensity score. Figures A.2 and A.3 in Appendix A show that the balance of scores between winning and losing startups greatly improves in the matched sample for stages 1 and 3, respectively. Table 1.6 shows the results of the PSM on the sample of ICT startups and the sample of science-based startups, which are in line with our main results. Moreover, the estimated treatment effects are very close to those estimated using the alternative identification strategies.

Table 1.6: PSM: Treatment effect of winning stage 1 and 3

	Science		Internet and Software	
	(1) Survival 2	(2) Rounds 2	(3) Survival 2	(4) Rounds 2
Panel A: Certification				
Wins Stage 1	0.395***	0.410***	0.131	-0.0307
	(0.132)	(0.135)	(0.185)	(0.127)
Panel B: Cash prize				
Wins Stage 3	0.0153	0.278	0.255***	0.407
	(0.0130)	(0.220)	(0.0730)	(0.297)
Observations stage 1	365	502	306	391
Observations stage 3	131	162	47	54

Note: Table presents the average treatment effect from a propensity score matching for stages 1, and 3. The level of observation is a startup in a competition round. Errors are clustered by competition round. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

In a nutshell, we have shown that our findings are robust to alternative specifications and identification strategies. The certification effect has a significant and positive effect on the future performances of science-based startups but does not affect ICT startups. Conversely, only ICT startups benefit from the cash prize.

1.8 Conclusion

The growing popularity of venture competitions has been accompanied by a rise in scholarly interest. The literature has already established that venture competitions can have a significant impact on startups' fate. However, there is still no consensus on the exact nature of the causal effect of venture competitions on startup performances. Some scholars argue that venture competitions are useful because the cash prize they offer alleviates startups' financial constraints (McKenzie, 2017). Others argue that what matters most is that they offer an informative certification of startup quality (Howell, 2020). In this paper, we reconcile these results and argue that, under the right circumstances, both the certification and cash prize can improve startups' business success. We propose a model of the impact of venture competitions on participating startups. We then test the model's predictions on a sample of 987 startups that have participated in Venture Kick, the leading Swiss venture competition. In line with previous findings, we find that venture competitions offer substantial benefits to participating startups. However, these benefits are not distributed equally among participants.

First, we find that the certification provided by Venture Kick's first stage has a long-lasting impact on science-based startups' survivability and ability to attract external funding. It helps high-quality startups secure financing and accelerates the termination of low-quality startups.

However, we find no such evidence for ICT startups. This differential effect appears to stem from differences in the reliability of judges' ratings. As Venture Kick's judges are less able to objectively assess the quality of ICT startups than science-based startups, their certification is less informative to both entrepreneurs and investors.

Second, we find that winning Venture Kick's third stage, and its $100,000 cash prize, has no impact on science-based startups but improves ICT startups' survivability. However, the cash prize only increases ICT startups' short-term survival and never impacts startups' ability to attract external investors. This finding is evidence that the cash prize's main role is to extend startups' runway. As a result, the cash prize effect depends on how much extra runway it can buy. ICT startups, with lower running costs, can better take advantage of Venture Kick's cash prize than the more capital-intensive science-based startups. More generally, this observation implies that we can expect the cash prize to have a greater effect in contexts where it affords a longer runway, either because it is generous in size or because local costs are low. This finding would explain why, on the one hand, McKenzie (2017) finds that young firms in Nigeria, a country with relatively low labor costs, benefit significantly from receiving a $50,000 cash prize and, on the other hand, Howell (2020) finds that the effect of cash prizes on U.S. startups is relatively small.

Although venture competitions' cash prize and certification have taken the limelight, venture competitions also offer other value-added activities. Indeed, venture competitions usually come with valuable mentoring and networking opportunities, which can prove crucial to the development of budding entrepreneurs (Timmons and Bygrave, 1986). Besides, tournament theory posits that competitions can be useful because they provide an incentive for participants to put some effort into improving their startup (Connelly et al., 2014). Finally, competitions can help entrepreneurs improve their startups' ideas and products, either through the experience they gather in a learning-by-doing framework (Arrow, 1971; Minniti and Bygrave, 2001) or thanks to the judges' feedback. We know little about the causal effect of these value-added activities, which presents avenues for further research.

The findings presented in this paper bear implications for the design of entrepreneurial programs. The first one is that a competition's certification can have a broad impact, both on startups' survivability and funding prospects. Moreover, this impact can be long-lasting. However, for the certification effect to materialize, organizers need to ensure that their judges are able to accurately assess the startups' potential. Only then will entrepreneurs and investors give credence and value to the certification. The second implication is that the cash prizes' role might be narrower than what organizers would hope for, only extending startups' runway without greatly impacting their long-term prospects. Organizers should therefore consider how much extra runway their cash prize can buy when designing venture competitions.

2 The Role of Venture Capital and Governments in Clean Energy: Lessons from the First Cleantech Bubble

This chapter is written in collaboration with David Popp.[1]

Abstract:

After a boom and bust cycle in the early 2010s, venture capital (VC) investments are, once again, flowing towards green businesses. In this paper, we use Crunchbase data on 250,000 U.S. startups founded between 2000 and 2020 to better understand why VC initially did not prove successful in funding new clean energy technologies. Both lackluster demand and a lower potential for outsized returns make clean energy firms less attractive to VC than startups in ICT or biotech. However, we find no clear evidence that characteristics such as high-capital intensity or long development timeframe are behind the lack of success of VC in clean energy. While public sector investments can help attract VC investment, the ultimate success rate of firms receiving public funding remains small. Thus, simulating demand will have a greater impact on clean energy innovation than investing in startups that will then struggle through the "valley of death". Only with demand-side policies in place should governments try to plug funding gaps by targeting clean energy startups with low potential for outsized returns that will continue to find it hard to attract private capital.

[1] The doctoral candidate is the main contributor to this paper.

2.1 Introduction

At the start of the last decade, venture capital (VC) experienced a boom and bust in clean energy technologies. From 2005 to 2008, the share of venture capital investments going to clean energy technologies more than tripled before falling in subsequent years. Today, once more, billions of dollars are pouring into green businesses. We argue that this renewed interest in clean technologies warrants further study of the initial failure of VC in this sector. Years have passed since the original assessments of the cleantech bust (Nanda et al., 2015; Gaddy et al., 2017), and the greater data availability today allows us both to draw a more complete picture and to contemplate the possibility of success for this new wave of green venture capital investment.

Moreover, providing support for green startups remains a popular policy tool. For example, in 2016, California created the Energy Innovation Ecosystem to support clean technology ventures, which includes seed funding through California Sustainable Energy Entrepreneur Development Initiative (CalSEED) awards. Similarly, the NY Green Bank, founded in 2014, leverages private sector financing to increase clean energy investments (Popp, 2020). The potential success of these initiatives depends on why early venture capital efforts failed. If issues such as high capital intensity or long delays between initial investment and commercialization create additional barriers for clean energy to raise funds, government financing can help bridge that gap. But if these investments are not successful because of a lack of demand for clean energy, government investments in these startups will be no more successful than private-sector VC investments.

To study the causes of VC's failure in cleantech, and in particular in clean energy, we obtained data on 251,108 companies launched in the United States between 2000 and 2020 from Crunchbase. The comprehensive coverage of Crunchbase's data allows us to compare venture capitalists' (VCs) performances in clean energy and electric vehicles (EVs) to their performances in other sectors with dynamic startups ecosystem like information and communications technology (ICT) or biotech. Using this data, we explore three factors that could explain the failure of VC in cleantech: financial constraints specific to funding clean energy, relative weak demand for clean energy products and limited potential for outsized returns.

Varying levels of capital intensity and observed VC returns allow us to analyze the role of financial constraints. Our results do not suggest that these financing constraints are the main factor behind the lack of success of VC in clean energy. For instance, while Bumpus and Comello (2017) argue that digital and modular cleantech startups should be better able to attract early-stage capital providers than capital-intensive energy hardware, recent data shows that digital energy firms have been no more successful than other energy startups.

Instead, we argue that weak demand for clean energy technology is the main reason for poor venture capital performance. To demonstrate the importance of expected demand, we implement a differences-in-difference estimation (DiD) centered around the passing of Ted Kennedy and the unexpected election of a Republican, Scott Brown, to replace him in

2010. Brown's victory in a special election in January 2010 made passing a comprehensive climate bill very difficult (Goldenberg, 2010), providing us with an exogenous negative shock to demand expectations for clean technologies. We find that expectations of weaker demand have a significant impact on VCs' willingness to fund clean energy startups. Finally, we show that clean energy firms have a lower potential for outsized returns than startups in ICT or biotech, making them less appealing to VCs.

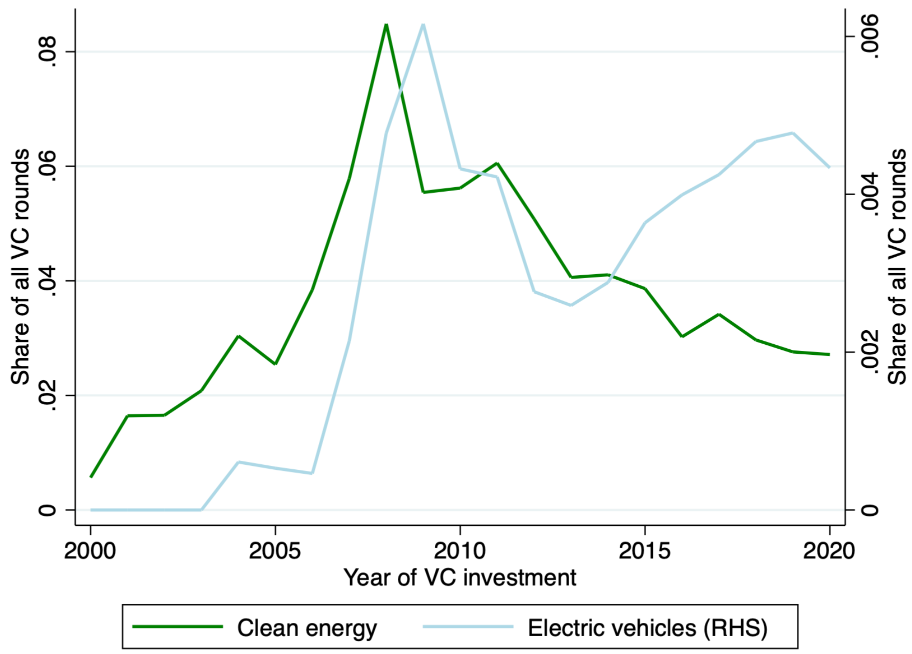
Our findings suggest that if governments want to support early-stage investments in clean energy, they should implement policies, such as carbon pricing, that stimulate widespread private demand for clean energy products. However, implementing such policies in the United States is an arduous political task, as demonstrated by the failure of a cap-and-trade bill in 2009-2010 when Democrats controlled the Presidency and both houses of Congress. The recent inclusion of funding for green technologies in the Infrastructure Investment and Jobs Act attests that spending money to support clean energy is more palatable politically. Thus, we also look at the role of the government as an active provider of funding to young companies. We show that in early stages, public investors provide small-sized grants that help startups prove the viability of their project and secure Series A funding, but that these firms are no likelier to have long-term success than other firms. In later stages, governments provide significantly larger sums aimed at helping startups expand and scale their business. When providing later-stage funding, we find that public investors have not fared worse than their private-sector counterparts, and may improve chances for exit if one assumes that public funding goes to startups less likely to attract late round private funding on their own. Here, our comprehensive data including companies in multiple sectors provides additional context compared to earlier studies focusing solely on public investments in the energy sector (e.g. Howell, 2017; Goldstein et al., 2020). Since venture capital investments in clean energy as a whole perform much worse than in other sectors, simply matching private sector performance in the clean energy sector is not enough. Public investments alone cannot plug all the funding gaps in cleantech innovation. We, therefore, argue that a positive demand shock is needed for private investors to be consistently successful when funding clean energy startups. Only then can governments use targeted public sector investments to address any remaining gaps in the valleys of death for cleantech innovation, such as by providing financial support for firms developing products where the potential for outsized returns is limited.

The paper is organized as follows. The next section provides background information on venture capital, its failure during the cleantech boom and governments' role in addressing funding gaps. Then, Section 2.3 presents our data. In Section 2.4 we develop our analysis of the underlying causes of the failure of VC in cleantech. Section 2.5 details and evaluates the role of governments in addressing funding gaps. Finally, Section 2.6 discusses our findings and their implication.

2.2 Background

The years 2005 to 2008 saw rapidly growing interest in clean energy from venture capitalists, policy-makers and the media. However, clean energy startups proved to be an unprofitable experiment; Gaddy et al. (2017) estimate that less than half of the over $25 billion provided to cleantech startups from 2006 to 2011 was returned; as a result, the cleantech boom went bust as VC funding dried up. Figure 2.1 shows that while more than 8 percent of VC rounds (Series A to J) reported in Crunchbase went to clean energy firms in 2008, this figure dropped to around 3 percent between 2016 and 2020. While most clean energy VC activity remains low, Figure 2.1 also shows renewed investor interest in EVs, which experienced a similar boom from 2006-2009 and are now returning to those previous levels of investment.

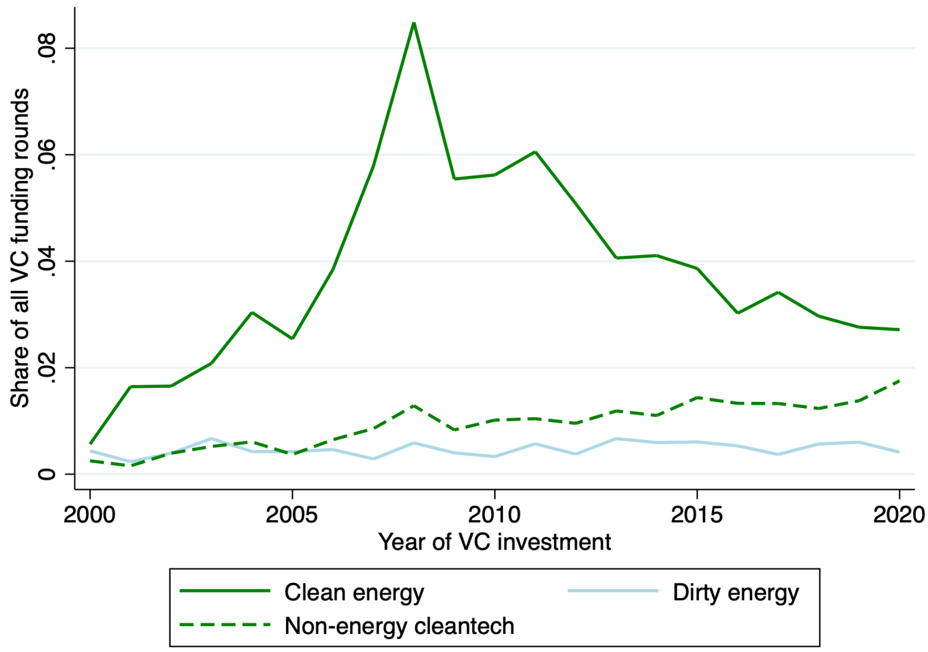
Figure 2.1: The cleantech boom and bust



Note: This figure shows the share of all VC rounds (Series A to Series J) going to clean energy (LHS) and electric vehicles (RHS).

Moreover, while dubbed the cleantech bubble, this bubble was concentrated in sectors related to clean energy, such as clean energy production, energy efficiency or EVs. Figure 2.2 shows that non-energy cleantech (e.g., sustainable farming, recycling or carbon capture) has been growing steadily since the start of the millennium without displaying a boom and bust pattern. Additionally, the bubble did not affect all energy firms equally. The share of VC funding to dirty energy (i.e., fossil fuels) has indeed remained relatively constant over the past two decades.

Figure 2.2: The bubble was concentrated in clean energy



Note: This figure shows the share of all VC rounds (Series A to Series J) going to clean energy, dirty energy and non-energy cleantech.

Failure of VC in energy

Several studies done in the wake of the cleantech boom and bust provide explanations for VCs' poor performance (Nanda et al., 2015; Gaddy et al., 2017; Saha and Muro, 2017). First, as a seller of a commodity, clean energy startups compete in markets with thin margins. The electricity produced by sun or wind is indistinguishable from coal-produced electricity, making product differentiation difficult. On top of that, in sectors like biotech, incumbent firms are willing to buy promising startups before commercial viability has been proven. Big energy firms and utilities have been less active in acquiring promising clean energy ventures, which makes investing at early stages less appealing. Finally, and most importantly, these papers suggest that high capital intensity, as well as long-lasting technological and commercialization uncertainty, are to blame for VCs' failure in clean energy. Overall, they argue that the VC model seems poorly suited for clean technologies (Lerner and Nanda, 2020).

Venture capitalists (VCs) are the primary source of early-stage funding targeted at young companies with the potential for rapid growth and high returns (Metrick and Yasuda, 2021). In addition to funding their internal growth, VCs monitor and mentor the startups in their portfolio. VC is structured as a fund whose manager (or general partner – GP) invests the capital of investors (limited partners – LPs) in a portfolio of companies. After an initial investment period of around 5 years, the GP has 5 to 7 years to return capital to the LPs (Kaplan and

Schoar, 2005). The goal of VCs is to exit their investments via an acquisition or an initial public offering (IPO) within the lifespan of the VC fund.

VCs invest at different stages of a startup lifecycle, providing capital to entrepreneurs in a stepwise manner (Neher, 1999). This process, called staged financing or VC staging, is a way to mitigate the agency problem. Rather than providing all the necessary capital to exit upfront, the VC investor retains the option to abandon a startup if it fails to meet stage targets (Tian, 2011; Nanda and Kropf, 2016). VCs start to become involved at the seed stage alongside angel investors, where a small amount of funding (i.e., usually less than $1M according to Crunchbase) is provided to prove a concept. Then during Series A, the first of the typical VC rounds, VCs offer more than $7M on average to young companies with a track record that illustrates the potential of the idea or technology and a solid strategy to transform this idea into a profitable business. As the company matures and proves its potential, the rounds become increasingly generous (Series B, C, D, etc.). Most papers that study the investments decisions of VC investors focus on Series A as it corresponds to the stage where the initial investment decision is made and then follow these Series A investments to see their eventual outcome (Nanda and Rhodes-Kropf, 2013; Gaddy et al., 2017; Ewens et al., 2018).

Investing in early-stage companies offers the opportunity for huge gains but also staggering losses. The expectation is that within a portfolio of 10 to 20 companies, a few will succeed, some will break even, and most will fail (Gaddy et al., 2017). The few successes need to be able to offer sufficient returns to compensate for all the failures. Therefore, the ideal investment will be easily and rapidly scalable with the potential to return 10 to 100 times the amount invested within 5 to 7 years. Because of that, scholars have argued that clean energy startups' high-capital intensity and long-lasting uncertainty around the technology and its commercial viability simply do not fit the VC model. The high capital intensity reduces the number of projects that a given VC fund can include in its portfolio, thereby increasing risk. Moreover, VCs' staging approach is at odds with clean energy startups' long-lasting technological and commercialization uncertainty.

In this paper, we explore two other reasons for the failure of VC in clean energy. The first is that demand for clean goods is simply too low because environmental externalities are not priced in. The peak in VC investment in clean energy & EVs, 2009, coincides with the failure of the cap-and-trade bill in the US Congress and the disappointing Copenhagen Climate Conference (COP 15). Investors had to subsequently reassess their expectations of future policy support for clean technologies. Second, we document the inability of clean energy firms to earn oversized profits. As noted earlier, clean energy companies find it difficult to differentiate their products and boost their market power. As a result, compared to startups in sectors like ICT or biotech, they struggle to earn the 10 to 100 times returns that VC investors look for.

Valleys of deaths

Funding immature clean technologies is, therefore, fraught with challenges. As a result, clean energy startups face two valleys of death, i.e., periods during which they might fail because of a lack of funding. The first valley of death, common to all startups irrespective of technology, is a technological valley of death that occurs between the basic research or idea and the viable prototype (Hudson and Khazragui, 2013). Startups can cross this valley of death if they find investors, usually VCs, to fund their early development. Receiving Series A is, therefore, a sign that a startup has crossed this first valley of death. At this point, startups in sectors such as ICT will have a product (e.g., software) that is easy and cheap to mass-produce (Hartley and Medlock, 2017). Their success or failure will then mostly depend on whether there is a demand for this product but not because of a lack of funding.

In other sectors, startups that have proven that their technology is viable at the prototype-phase still need to prove that their product is viable on a commercial scale. At this stage, startups potentially enter the second commercialization valley of death. The financing required to commercialize is often too high for early-stage VCs and traditional lenders in the private sector are reluctant to fund technologies that have not been proven at scale (Frank et al., 1996). Notably, energy innovations are often more capital intensive and take longer to commercialize than other sectors (Popp, 2017). However, other industries share similar attributes. For biotech firms, the average cost of bringing a drug from concept to market is estimated to be above $1B and can take approximately 11 to 12 years (Hudson and Khazragui, 2013). However, biotech benefits from having incumbent firms willing to buy promising startups before they have proven their commercial viability (Nanda et al., 2015). In contrast, the second valley of death remains an important hurdle for the development of new clean technologies (Popp et al., 2020). Its presence implies that even after having convinced Series A investors to fund their early project, viable startups can get stranded for lack of funding.

The role of public investors

Both the first technological valley of death and the second commercialization valley of death in clean technologies raise the question of the role of governments in providing funding. Many governments programs provide support to very early-stage startups, which allows them to prove their technological viability and attract Series A funding (e.g., CALSEED). In other words, they help startups bridge the first technological valley of death. Tellingly, Howell (2017) finds that the US Department of Energy (DOE) Small Business Innovation Research (SBIR) R&D grants are helpful because they fund technology prototyping. Moreover, she shows that this early-stage award increases the probability of securing subsequent venture capital: in other words, it increases the probability of bridging the first valley of death. However, early public support does not seem to address the second valley of death for cleantech startups. Goldstein et al. (2020) find that receiving an early-stage award from the US Advanced Research Projects Agency-Energy (ARPA-E) does not significantly increase the probability of exiting compared to comparable but non-participating cleantech startups.

Governments can also intervene after Series A rounds to help bridge the commercialization valley of death. Both Mowery et al. (2010) and Weyant (2011) argue that government financing can help new energy technologies overcome roadblocks to commercialization. For instance, significant energy innovations typically have disproportionately large capital expenses, leaving a role for collaboration with the public sector to support initial project development and demonstration projects. Other scholars argue that, given governments poor track record in the governance of large demonstration projects, this situation is a dilemma between market failure (valley of death) and government failure (Nemet et al., 2018). An example of a program aimed at bridging the commercialization valley of death is the DOE Loan Guarantee Program, part of President Obama's stimulus program. The DOE's Loan Programs Office (LPO) has over 30 projects in 18 states and has enabled over 50 billion dollars in private sector investment in clean energy technologies (Congress, 2017).[2] While this program is remembered for the failure of Solyndra, another beneficiary of this loan program was Tesla, who received $465M in direct loan, repaid it in full in 2013 and is today the most valuable cleantech company in the world. Moreover, there is little evidence that the program performed worse than its private-sector counterpart; if anything, it was relatively successful (Lewis, 2018).[3]

Public investors can, therefore, directly fund startups to help them bridge the valleys of death. However, their usefulness depends on the underlying causes of the failure of private investors in cleantech. If clean energy companies' unattractive characteristics (i.e., capital intensity and long development timeframe) are to blame, then patient public capital can indeed help. If a lack of demand for the eventual green product is the culprit, public investments cannot be expected to perform much better than private investors and will not solve the funding gaps alone. Understanding the causes of the failure of VC in clean energy is, therefore, key to guiding public policies.

2.3 Data

We gather firm-level data on startups and funding activity from Crunchbase, a provider of business data for private and public companies. Crunchbase provides detailed information on companies, such as their founding date, headquarters' location at the postal code level, industry classification and founder characteristics (i.e., previous entrepreneurial experience, degrees and gender). Most importantly, they provide detailed information on funding rounds as well as exits (i.e., date, amount, investment/exit type and actors involved). This information includes the type of organization making each investment, such as a government office, venture capital, or an investment bank.

Crunchbase collects its data through companies and investors' self-submissions, the work of their own data analysts and by using AI and machine learning. As a result, there may be

[2]The entire portfolio is available on the DoE's website: https://www.energy.gov/lpo/portfolio-projects.

[3]This success is reflected in a quote from Michael Lewis' book, The Fifth Risk: "*The DOE had built a loan portfolio that, as MacWilliams put it, "JPMorgan would have been happy to own." The whole point was to take big risks the market would not take, and they were making money!*"

selection concerns, for instance, as more innovative companies are more likely to appear in the data. Contemporaneous and past coverage increase over time as new data is made available and as more companies use the platform (Dalle et al., 2017). In the final one or two years, coverage is more partial given the time lags to collect recent data. Furthermore, some firms may misleadingly indicate that they operate in a particular sector for self-promotion purposes to attract more funding, as sector categories are not cross-checked against traditional sectoral classifications. While we do not explicitly model selection into Crunchbase, our regression analysis uses year and sector fixed effects to partially control for such selection issues. For instance, year fixed effects control for changes in the completeness of Crunchbase coverage over time

Our dataset

To conduct our analysis, we link nine different datasets provided by Crunchbase.[4] Beginning with the whole cross-section of 1,382,795 organizations registered in Crunchbase, we keep organizations located in the United States that define themselves as companies (as opposed to an investor or a school) and who launched between 2000 and 2020.[5] This reduces our sample down to 251,108 companies. We then match these companies to Crunchbase's funding-round level dataset. Only 74,509 companies (around 30%) of these companies have at least one recorded deal. These 74,509 companies secured a total of 165,241 individual funding rounds.[6] While we use all the 28 funding types registered by Crunchbase in our analysis, we focus mainly on venture capital rounds (Series A, B, ...) and, to a lesser degree, on seed rounds.[7] Additionally, using Crunchbase investors' dataset, we can determine whether a VC or a public investor is involved in each funding round.[8] Table 2.1 provides a basic description of the investment types as well as the involvement of public and VC investors in our matched 165,215 funding rounds.

[4]We use the organizations, acquisitions, ipos, funding_rounds, investments, investors, jobs, degree and people datasets.

[5]We downloaded the data using an Academic Research Access on the 25th of May 2021.

[6]One caveat is that the funding amount is not disclosed in 25% of the observations. As we do not make an extensive use of funding amounts and rather focus on funding dummies this does not significantly impact our analysis. The only time we use amounts is when looking at the returns for investors in Series A rounds, for which only 14% of the amounts are undisclosed.

[7]Crunchbase provides a, not entirely exhaustive, overview of its funding types: https://support.crunchbase.com/hc/en-us/articles/115010458467-Glossary-of-Funding-Types

[8]We are able to identify the investors in 65% of the 165,241 funding rounds. This number rises to 85% during the Series A rounds. To improve the data coverage in Table 2.9, we assume that all Series A to J rounds include a VC because, when the data is available, we find that VCs are involved in Series A to J rounds 85% to 93% of the time.

Table 2.1: Descriptive statistics of 165,215 funding rounds

	All	Seed	Series A	Series B	Series C	Series D
Share of all rounds		29.9%	11.8%	6.4%	3.2%	1.5%
Share that involves public investors	4.7%	4.4%	2.8%	2.7%	3.2%	3.9%
Share that involves VC investors	63.3%	60.8%	84.8%	90.2%	91.5%	92.0%

Note: The first row is the share of all rounds going to each funding type. The second and third row are the share of each funding type that involves public or VC investors.

To measure the success of startups, we match our 251,108 companies to 3,044 initial public offerings (IPOs) and 22,570 acquisitions. However, data availability for the values of these exits is limited. Only 15% of acquisitions and 27% of IPOs disclose the amount. Thus, in most of our analysis, we use binary exit variables rather than exit values. When we do use exit amounts, we only do so to complement our analysis and only for clean energy & EVs, ICT or biotech startups that received Series A investment, for which data availability is higher (amounts are disclosed for 26% of acquisitions and 68% of IPOs). All funding and exit amounts are converted to real 2020 dollars using the US CPI from the Bureau of Labor Statistics (BLS).

Sector Analysis

Crunchbase's comprehensive coverage allows us to look at multiple sectors, thereby comparing the performances of investors in clean energy to their performances in other dynamic sectors. We focus only on the 149,358 startups operating in our three sectors of interest: clean energy & EVs, ICT and biotech. Table 2.2 provides detailed descriptions of each sector. We include ICT companies because they provide a relevant benchmark to the success of clean energy firms, as ICT is the sector with the most active startup ecosystem. We also include biotech firms because they provide another interesting comparison group that has attracted a lot of VC investments despite its high capital needs. Our clean energy & EVs, ICT and biotech categories account for a disproportionate share of funding activity in the United States. They represent 59% of the startups in our dataset but are responsible for 75% of the funding recorded in Crunchbase.

As shown in Table 2.2, clean energy & EVs includes companies focused on supply and demand of clean energy. Supply includes companies working on technologies related directly to clean energy production, such as solar and wind energy, or those working indirectly through grid management. Demand-side technologies include energy efficiency and electric vehicles.[9] Such companies played a central role in the VC boom and bust of the early 2010s. Among clean energy companies, we differentiate between the companies that develop energy technology with some digital or ICT elements and those that do not. Examples of digital energy startups include companies that develop smart sensors or software that optimize battery-related

[9] It is important to note that we categorize as clean energy all the energy startups that are not involved in fossil fuel technologies. We also include in our clean energy category the startups involved in both fossil fuel and one of the clean categories in Table 2.2

processes. Popp et al. (2020) document increasing cross-fertilization between clean energy and digital technologies. As we show in this paper, digital energy startups are less capital-intensive, which could make them more attractive targets for venture capital (Gaddy et al., 2017). As electric vehicles were a small portion of the previous clean energy VC boom but are attracting growing attention from investors (see Figure 2.1), we treat electric vehicles as a separate category in our analysis. The top panel of Table 2.2 shows the composition of relevant industry categories, and its bottom panel displays the criteria for each company's placement into our five mutually exclusive categories (i.e., non-digital clean energy, digital clean energy, EVs, ICT and biotech).

Table 2.2: Firm classifications and descriptions

Firm classification	Crunchbase industry categories	Observations
All energy		6395
Renewables	Clean Energy, Renewable Energy, Storage, Solar, Wind	3112
Fossil fuel	Fossil Fuels, Fuel Cell, Oil and Gas, Fuel	1577
Grid Management	Electricity Distribution, Energy Management, Power Grid	795
Energy Efficiency	Energy Efficiency	569
Other Energy	All other energy types, including biomass and biofuel	1155
Electric Vehicles	Electric Vehicle	270
ICT	Apps, AI, Data, IT, Internet, Telecom, Mobile, Platforms, Software, E-Commerce, Online Auctions	138544
Biotech	Bioinformatics, Biometrics, Biopharma, Biotechnology, Genetics, Life Science, Neuroscience, Quantified Self	9582
Category	Description	Observations
Clean energy & EVs	Clean energy + Electric vehicles	5222
Clean energy	All energy excluding fossil fuel companies	4952
Non-digital clean energy	Clean energy that is not also ICT	*3829*
Digital clean energy	Clean energy also ICT	*1123*
Electric vehicles	All EV	*270*
ICT	ICT but not clean energy & EVs or biotech	134810
Biotech	Biotech but not clean energy & EVs	9326
Total		149358

Note: This table's top panel display all the industries of interest to our study, with the constituting sub-categories. The bottom panel defines the exact, mutually exclusive, industry categories that we use in our analysis.

Measures of success

As we are interested in studying the performances of early-stage investors in clean energy, we focus on the success of investors involved during Series A rounds.[10] The success of VC investors can be measured by whether they have been able to exit their early-stage investments through an initial public offering (IPO) or an acquisition by another company or later-stage investor (Metrick and Yasuda, 2021). We, therefore, use startups' exit events to draw conclusions about the performances of VCs across sectors. In addition to exit dummies, we also develop more granular measures of VC investment return (i.e., a measure of the success of the funded startup). We indeed want to be able to differentiate between the startup that exit while returning 2x invested amount to their Series A investors from those that are home run successes and yield a 100x return to their early-stage investors.

According to Gompers et al. (2020), most VCs use the ratio of return over invested capital to analyze their investments, rather than more complex measures that discount future returns (e.g., IRR or NPV). We, therefore, develop the following cash-on-cash (CoC) multiple metric to analyze and compare the performances of VC across sectors:

$$CoC_i = \frac{\text{Value at exit}_i * \text{VC stake}}{\text{Series A amount}_i}$$

Value at exit is either the valuation at IPO or the acquisition price.[11] If startup i does not exit, its value at exit is set to 0. The *Series A amount* corresponds to the total value of the Series A rounds raised by startup i. Finally, the stake corresponds to the ownership stake that Series A investors have in the startup i. Unfortunately, ownership stakes can vary substantially and are typically not disclosed. Therefore, in line with Gaddy et al. (2017), we assume that VCs stake in Series A is 12%.

While the CoC is a valuable metric that allows us to compare investments on an equal footing (i.e., a startup that require more capital in Series A will need a higher value at exit to compensate), it also has some drawbacks. The first is that we need to make an assumption about the ownership stake. The second is that we do not have the amount for 14% of the Series A rounds. To address these, we also look at the exit value (i.e., IPO valuation or acquisition price) by themselves, a simpler but more widely available metric. Finally, both our CoC and exit value variables suffer from the fact that exit amounts are only disclosed for 26% of the acquisitions and 68% of the IPOs of startups with Series A. To address this final problem, we create four binary variables that indicate whether startup i yielded at least a 5x returns, 10x returns, an exit above $1 billion or above $5 billion. We then assume that if the exit amount is not disclosed, the startup did not yield at least a 5x return or an exit above $1 billion. We can make this assumption because undisclosed exits are often a sign that an investment did not return capital to its investors (Gaddy et al., 2017). This allows us to increase our data

[10] In our dataset, companies secure their first Series A round on average three years after their launch.
[11] If both events happen, we use the IPO value as it is the event that occurs first in 88% of cases, and is therefore when investors will cash in their returns.

availability significantly while keeping some of the granularity in returns.

Dealing with living-dead companies

While identifying the successful investments is relatively straightforward, we also need to identify the failed investments, which are harder to determine with complete certainty. Indeed, 11% of clean energy, EVs, digital and biotech companies have either been acquired or gone public but not all of the remaining companies are failures. Only a minority (5%) have been officially "closed" and the rest (84%) are still categorized as operating. It can be challenging to determine whether operating companies will end up as successes or failures. Some of them may be growing, raising new funding and will eventually exit. Others will continue operating without any prospect of exit, becoming so-called "living dead" companies (Ruhnka et al., 1992). These living-dead companies can be considered failures, as VCs' objective is for their startups to produce a large exit within 5-7 years of the initial investment (Metrick and Yasuda, 2021). To separate these "living dead" from the active companies that are still growing, we use the fact that, in our dataset, around 75% of acquisitions and 85% of IPOs occur within three years of their previous funding round. We therefore assume that a period of inactivity of at least three years is a sign of being "living dead". Based on that, we classify operating startups into two groups. Those that have been launched or funded in the three years before the 1st of May 2021 but have not yet exited are categorized as "still active". They are excluded from any analysis looking at exit and the associated returns (i.e., when the dependent variable is an exit dummy, our CoC metric or the exit value) as we do not know what their eventual outcome will be (i.e., exit or failure). The companies that have not received or been funded in the three years before the 1st of May 2021 are categorized as failures (i.e., they have not exited) because even though they might technically still operate, they will most likely not return any capital to their investors. In the end, in our sample of 149,358 companies, 10.8% have exited, 14.6% are still active and 74.6% are failures (i.e., they have not exited).

The time horizon of the regressions

Finally, in our regressions analyses we look at startups launched or funded between 2003 and 2016. We do not include the startups launched before 2003 because they are less affected by the cleantech boom and therefore less informative and we do not include those launched after 2016 because these companies simply have not yet had the time to exit and return money to their investors.

2.4 The underlying causes of VC's failure in clean energy

While several papers document the poor performance of early clean energy VC investments, understanding why these investments fail is important for understanding what role, if any, public investors can play in supporting clean energy startups. This depends on the extent to which financial constraints themselves hinder clean energy investments relative to a lack of demand. That is, are the valleys of death for clean energy startups really due to the particular characteristics of energy innovation making clean energy unattractive to VC investors, or is it simply a result of historically underpriced environmental externalities reducing demand for cleaner technology?

In this section, we present evidence that the lack of demand for clean energy is a crucial reason for VCs' poor performances and low interest in clean energy. Although we cannot definitively rule out that financial constraints for clean energy exist, we find no evidence that increasing financing for clean energy technologies will, on its own, be successful. Additionally, we document the low potential for outsized returns in clean energy for early-stage investors, which is not only a consequence of subdued demand but also, we argue, of the weak market power of energy companies.

2.4.1 The role of startups' characteristics and financing constraints

In this section, we test two hypotheses about the role of financial constraints in clean energy VC. First, after a first wave of, hard to fund, hardware-focused cleantech startups, investors and scholars have argued that the second wave of digital and modular cleantech startups should be better able to attract early-stage capital providers (Bumpus and Comello, 2017). This reasoning stems from the fact that ICT startups, in general, are an attractive investment for VCs because they have relatively low capital needs and the uncertainty about the viability of their product and the market demand can be resolved quickly (Lerner and Nanda, 2020). Supporting this claim, Gaddy et al. (2017) found that digital cleantech startups were the only ones to have posted positive returns for their Series A investors during the cleantech boom and bust. This leads to the following hypothesis:

Hypothesis 1: If clean energy startups' unappealing characteristics are the main reason behind VCs' failure in energy; the less capital-intensive digital energy startups should have both fared better and attracted more VC money since the bust than the rest of the clean energy sector.

Advocates of increased public funding for clean energy startups often argue that investors ignore clean energy because it is capital intensive and provides slow returns, so that clean energy firms have less access to VC than other sectors (Weyant, 2011). That is, clean energy firms face more liquidity constraints than other types of firms, so that likely profitable opportunities are unable to raise funds. To check for the presence of such financial constraints in the energy sector, consider the alternative. With equal access to capital in all sectors, investments should flow so that the expected return is equal across sectors. If liquidity constraints limit the flow of VC into clean energy, the limited funds that do flow to clean energy should go to the most promising clean energy firms. As such, the expected returns to investments in clean energy should be *highe*r than the expected returns in other sectors. On the other hand, if the flow of VC into clean energy is limited by lackluster demand, the expected returns should not be higher than in other sectors. While we can only observe actual, not expected, returns, we argue that evidence that investors are "leaving money on the table" would be evidence of significant financial constraints. Investors are "leaving money on the table" if the performances (i.e., realized returns) of the clean energy startups that do get funded by VCs are significantly higher than the performance of clean energy without VC funding, in comparison to other sectors. We call this differential between the performances of VC-funded and non-VC-funded startups the *VC premium*. Therefore:

Hypothesis 2: If financial constraints affect clean energy more than other sectors, investors will leave money on the table, leading to a higher VC premium for investments made in this sector.

Methodology & Results

We first use Crunchbase data to verify whether digital energy startups are indeed less capital intensive than other clean energy firms. To do so, we look at the money raised after Series A round by the startups that eventually exited, excluding the funds raised after their exit. While this is not a perfect measure of capital intensity, the amount raised is directly linked to the capital needed to grow a company to the point where it can exit. Table 2.3 shows that digital clean energy startups raise \$49.6M (2020 US dollars) after their Series A on average, nearly 50% less than the \$94.2M raised by non-digital clean energy startups.[12] At the same time, ICT companies have similar capital needs than digital energy companies, which is what we would expect. Biotech and non-digital energy companies have similarly high levels of capital intensity. Therefore, we can separate our sample into the capital-intensive companies (i.e., non-digital energy, EVs and biotech) and the less capital-intensive companies (i.e., digital energy and ICT).

Table 2.3: Average amount raised before exit by startups with Series A (2000-2020)

in 2020$ million	Non-digital en.	Digital energy	EVs	ICT	Biotech
Post-Series A Funding	94.2	49.6	216.1	50.1	80.4
Observations	78	46	9	3446	707

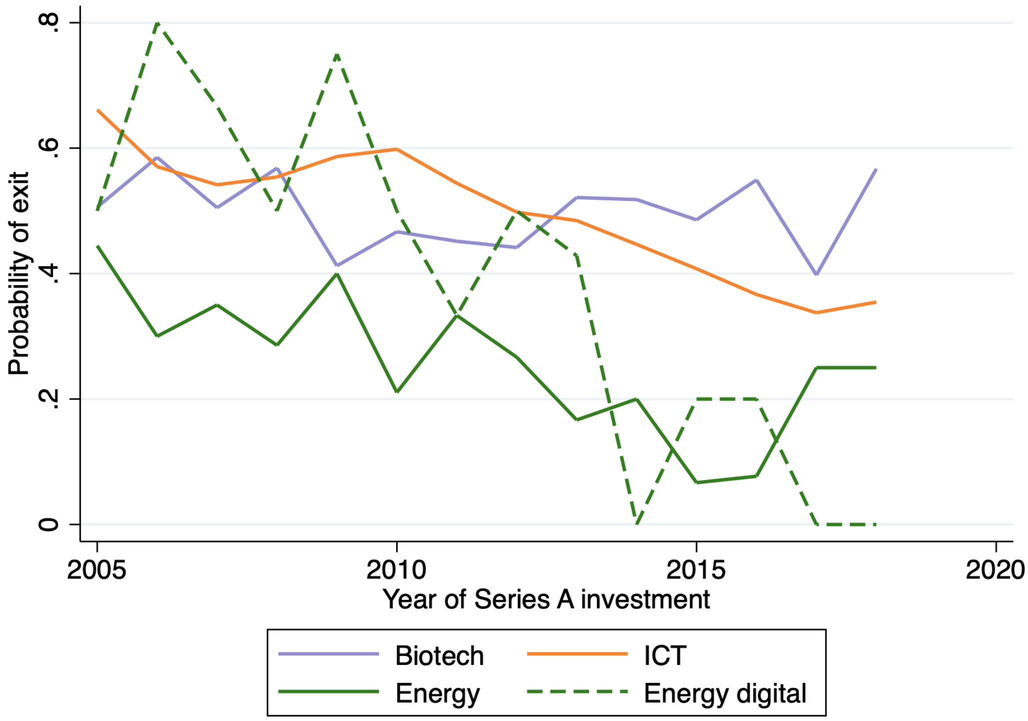
We first test Hypothesis 1 by examining VC performances and funding rates across sectors. As a measure of the performance of VC investments, we look at the exit rate of the startups that received Series A investments. Figure 2.3 shows that digital energy startups that received Series A up until 2013 had a similar probability of exit than VC-funded biotech and digital startups.[13]. This is evidence that during and directly after the cleantech boom and bust, digital

[12]EV companies raised \$216.1M on average but It is important to note that this average is only based on 9 observations, including Tesla and ChargePoint who have each raised more than \$680M.

[13]Note that in the figures below, we do not display results for EVs as the lower number of observations prevents us from providing stable yearly output. See Table 2.4 for EV results.

energy startups have outperformed more traditional energy startups, which is in line with the initial studies of the cleantech failure. However, Figure 2.3 also shows novel evidence that this initial outperformance did not stand the test of time. Digital energy startups funded after 2014 have indeed seen their exit rate fall to the low level prevalent in the rest of the energy sector.

Figure 2.3: The performance of VC-funded digital energy startups has gone down



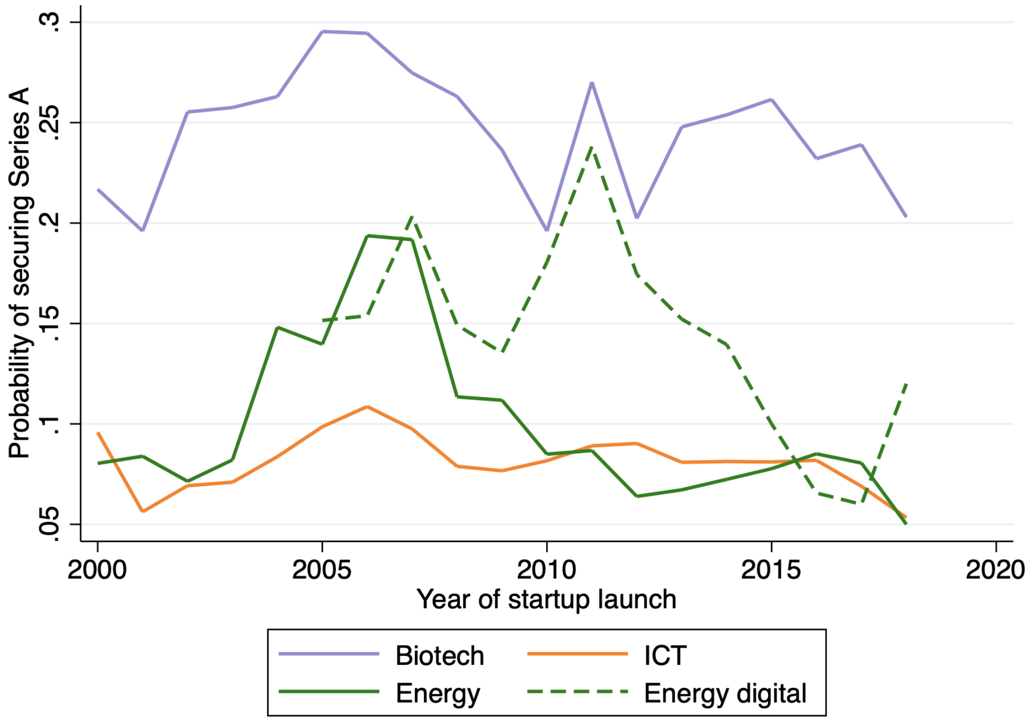
Note: This figure displays the share of Series A-funded startups that that managed to exit, based on the year they received the Series A round, by industry.

Our first finding contradicts the idea that digital energy, like ICT, is a good fit for VC. Next, we demonstrate that the worsening performances of digital energy companies has been accompanied by a fall in the willingness of VC to fund them. Figure 2.4 below shows that while energy startups launched after 2007 faced increasingly reluctant VCs, energy digital startups launched up until 2011 enjoyed buoyant VC interest. However, while 25% of energy digital startups launched in 2011 received Series A funding, only a little more than 5% of those launched in 2016 and 2017 did. The drop in VC's willingness to fund digital energy startups follows the fall in their exit probabilities to the level of the rest of the clean energy sector (see Figure 2.3 and Figure 2.5).[14] It is interesting to note that ICT startups have a relatively low funding rate. This stems from the fact that ICT startups are very common, representing

[14]Note that Figure 2.3 conditions startups based on the year of investment unlike Figures 2.4 and 2.5 that are based on the year of launch. This explains the lag between the two groups of Figures as, in our sample, startups receive their first Series A funding three years after their launch on average.

54% of the US startups on Crunchbase launched between 2000 and 2020, and thus face stiff competition for VC money.

Figure 2.4: VCs' willingness to fund energy digital startups has also fallen



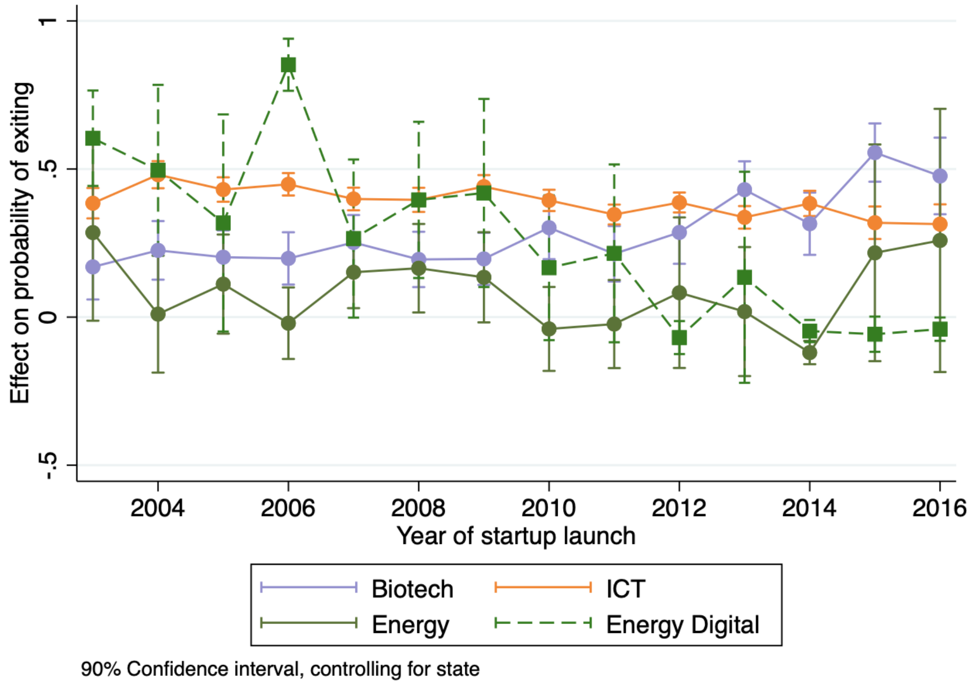
Note: This figures displays the share of all startups that manage to secure Series A funding, based on the year they were launched, by industry. Values are not displayed for energy digital startups launched before 2005 because of the low data availability (i.e., only between 3 to 8 startups were funded each year).

Turning now to Hypothesis 2, we test (a) whether startups supported by VCs have outperformed the startups that did not receive Series A funding and (b) whether this VC premium is higher than in other sectors. To calculate this VC premium over the years, we implement Equation 2.1 using OLS:

$$\begin{aligned} Exit_i = \beta_0 + \beta_1 VC_i + \beta_{2,ind} Industry_i + \beta_{3,ind} VC_i * Industry_i + \\ \beta_{4,year} VC_i * Launch\_year_i + \beta_{5,ind,year} Industry_i * Launch\_year_i \\ \beta_{6,ind,year} VC_i * Industry_i * Launch\_year_i + \eta_{launch\_year} + \eta_{state} + \epsilon_i \end{aligned} \tag{2.1}$$

Here, the unit of observation is startup *i*. The dependent variable *Exit* indicates whether the startup got acquired or had an IPO. *Industry* is a categorical variable that indicates whether startup *i* is clean energy, energy digital, EVs, digital or biotech. The excluded category is digital startups. The dummy variable *VC* captures whether a startup receives Series A financing. *Launch year* indicates the year in which the startup is founded, which controls for both for

Figure 2.5: The VC premium in digital energy has vanished



Note: This figure displays the marginal effect of receiving VC funding by industry and startup launch date, using Equation 2.1. Estimates are shown with their 90% confidence intervals.

common economic conditions and the changes in Crunchbase's coverage. Finally, we include state fixed effects. The year 2003 and the state of California are the reference categories.

Figure 2.5 below displays the VC premium by industry and year, i.e., Equation 2.1's $\beta_1 + \beta_{3,ind} + \beta_{4,year} + \beta_{6,ind,year}$. It shows that digital energy startups launched between 2005-2008 supported by VCs (i.e. those funded during the cleantech bubble) outperformed the non-VC supported digital energy startups and had a similar VC premium than the other (non-energy) ICT investments. However, this has stopped being true for startups launched after 2011. In the past decade, VCs in digital energy, like in non-digital energy, have not been able to pick startups that perform better than average. In comparison, the VC premium in biotech and ICT has never disappeared.

In order to show VCs premiums that are more easily comparable across sectors and for the pre and post-boom periods, we display the results from Equation 2.1 in Table 2.4. The interactions between receiving Series A venture capital and each sector allow us to test Hypothesis 2 formally. We estimate the VC premium by industry for all startups founded during the 2003 to 2016 period as well as during the pre-bust (2003-2009) and post-bust (2010-2016) eras.[15] Both

[15]As Figure 2.1 shows, the rapid growth in clean energy VC investments occurred during 2005-2008. After 2011, they have steadily fallen. Startups founded between 2003 and 2009 have therefore launched and likely received their first funding during the cleantech boom – as they receive their Series A 3 years after launch on average.

ICT (the excluded category) and biotech startups display a very significant VC premium; ICT startups that receive series A have an exit probability around 40 percentage points higher than those that do not.

Table 2.4: The relationship between Series A funding and eventual success

	General (2003 to 2016)			2003 to 2009	2010 to 2016
	(1) Exit	(2) Acquired	(3) IPO	(4) Exit	(5) Exit
Received Series A	0.3925***	0.3719***	0.0290***	0.4241***	0.3632***
	(0.0067)	(0.0067)	(0.0024)	(0.0096)	(0.0093)
Energy	0.0549***	0.0164***	0.0498***	0.0799***	0.0312***
	(0.0071)	(0.0062)	(0.0045)	(0.0112)	(0.0087)
Energy digital	0.0206*	0.0078	0.0120**	0.0354*	0.0098
	(0.0121)	(0.0112)	(0.0048)	(0.0213)	(0.0140)
Electric vehicles	0.0902**	0.0318	0.0668***	0.1375**	0.0639
	(0.0354)	(0.0297)	(0.0239)	(0.0671)	(0.0401)
Biotech	0.1073***	0.0481***	0.0679***	0.1336***	0.0862***
	(0.0061)	(0.0053)	(0.0039)	(0.0098)	(0.0076)
Series A x Energy	-0.2895***	-0.2626***	-0.0399***	-0.3162***	-0.3363***
	(0.0305)	(0.0290)	(0.0146)	(0.0367)	(0.0522)
Series A x Energy	-0.0727	-0.0339	-0.0242	0.0130	-0.2425***
digital	(0.0548)	(0.0548)	(0.0176)	(0.0712)	(0.0741)
Series A x EVs	-0.1288	-0.1633	0.0165	-0.1786	-0.1302
	(0.1258)	(0.1210)	(0.0813)	(0.1574)	(0.2281)
Series A x Biotech	-0.1156***	-0.2291***	0.1408***	-0.2138***	-0.0047
	(0.0173)	(0.0159)	(0.0136)	(0.0235)	(0.0259)
Constant	0.0368***	0.0310***	0.0082***	0.0379***	0.0783
	(0.0068)	(0.0066)	(0.0026)	(0.0068)	(0.0497)
Launch year FE	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes
Observations	104800	104800	104800	39868	64932
R^2	0.111	0.094	0.073	0.118	0.093
VC premiums					
Series A: Energy	0.103***	0.109***	-0.011	0.108***	0.027
Series A: Energy digital	0.320***	0.338***	0.005	0.437***	0.121
Series A: EVs	0.264**	0.209*	0.045	0.245	0.233
Series A: Biotech	0.277***	0.143***	0.170***	0.210***	0.359***

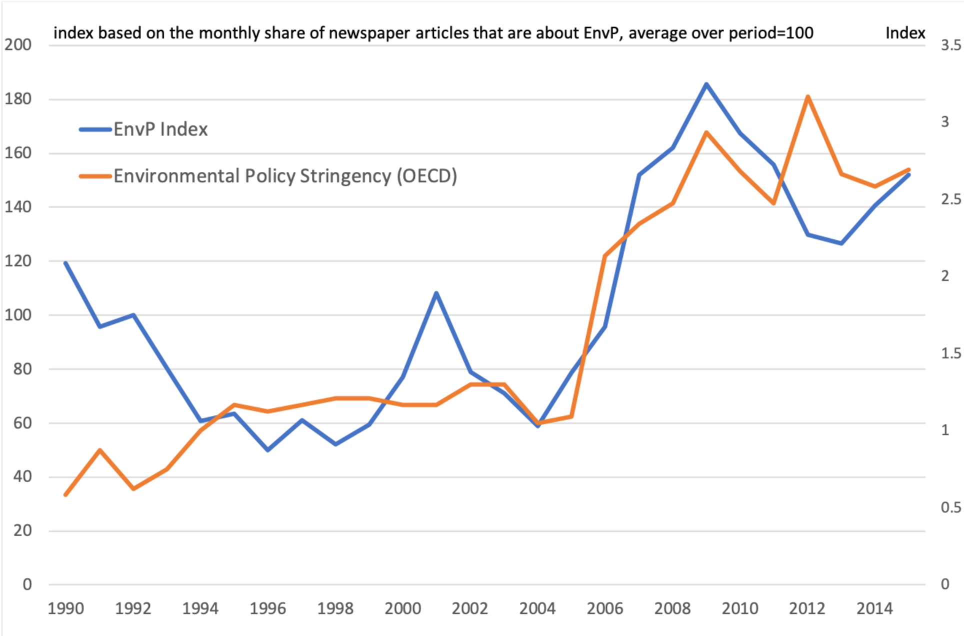
Note: This table presents results of OLS regressions. The sample includes clean energy & EVs, ICT and biotech startups launched from January 2003 to December 2016. The dependent variable is whether the startup exited in Columns (1), (4), and (5). In Column (2), the dependent variable is whether the startup got acquired and in Column (3) whether it went public. The bottom panel displays the VC premiums by industry, i.e., the sum of the *Received Series A* and *Series A x industry* coefficients and their statistical significance. Note that the *Received Series A* coefficient represents the VC premium for ICT startups. The level of observation is at the firm-level. Robust standard errors are in parentheses. Asterisks denote * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Overall, this section shows that less capital-intensive digital clean energy startups have quickly stopped outperforming the rest of the energy sector after the bust. What is more, EVs, a capital-intensive industry that is trying to normalize the use of a radical new technology, has seen a resurgence in VC funding (see Figure 2.1) and stable success (see Table 2.4). We also find that, since the bust, clean energy startups that receive VC funding do not perform any better than those that do not, unlike VC-funded startups in other sectors. The absence of VC premium in clean energy indicates the investors are not "leaving money on the table". These results cast doubt on the preeminence of startups characteristics, and the resulting financing constraints, as an explanation for the lack of success of VC in clean energy.

2.4.2 The role of expected demand for environmental goods

Having shown that the capital intensity of energy technology cannot, on its own, explain the poor performance of venture capital in the energy sector, we now show that changes in expected demand – caused by changes in policy support – do explain the investment patterns observed during the boom and bust period. Noailly et al. (2021) find that decreases in environmental policy stringency are associated with lower willingness from VC to fund cleantech startups. The first phase of the cleantech boom was accompanied by rising environmental regulations as well as expectations of increasingly stringent future environmental policies. The election of President Obama and a Democratic majority in Congress raised expectations for a national climate policy, which eventually became the Waxman-Markey bill (formally the American Clean Energy and Security Act) that passed the House in 2009. The bill proposed the first national cap-and-trade market for carbon emissions in the US. However, both the Waxman-Markey bill as well as the international Copenhagen Summit (COP15 in December 2009) ended in failure. This reduced VCs' expectations of future policy support and therefore of future returns in clean technologies. To illustrate that, Figure 2.6 shows the evolution of environmental policy in the United States using the OECD's environmental policy stringency index and Noailly et al. (2021)'s EnvP Index between 1990 and 2019. We can see that after 2009, the rapid increase in both environmental policy stringency and its media coverage stalled or even decreased.

Figure 2.6: Environmental policy stringency and coverage rose after 2004



Note: This figure displays the relationship between the Environmental Policy Stringency index of the OECD and the EnvP index constructed by (Noailly et al., 2021) at annual frequency.

Methodology & Results

To test whether changing expectations of future demand for clean energy products affect VC decisions, we implement an event study around the death of Senator Ted Kennedy in August 2009, a key vote in Obama's climate policy plans. His passing was followed by the upset victory of the Republican Scott Brown in January 2010 in a special election. Brown gave Republicans a 41st seat in the Senate, thereby robbing Democrats of a filibuster-proof majority. This meant that passing any comprehensive health or climate bill (e.g., the emissions trading Waxman-Markey Bill) became an impossible task (Davenport and Samuelsohn, 2010; Goldenberg, 2010). Moreover, this event affected demand-side climate policies particularly. First, because it was a demand-side policy (cap and trade) that Congress was trying to pass and that which Brown's election made impossible. And then, because Republicans are more comfortable with supply-side policies (e.g., the Advanced Research Projects Agency-Energy was created while George W. Bush was president) than demand-side policies that restrict what people can do (e.g., a carbon tax).

Comparing VC investments in clean energy and in a control group just before and after the unexpected victory of Senator Brown allows us to isolate the effect of a change in expected demand for clean goods on VCs' decisions. Before Senator Brown's victory, investors could expect future environmental policies to prop up clean energy startups. However, after Brown's victory, these expectations had to be readjusted. The clean energy startups funded after

Brown's election need to be able to survive in a policy environment that is now expected to be less supportive. Therefore, if demand plays a significant role, we expect Series A investors to become more demanding, increase their quality threshold and fund less clean energy startups after Brown's election. The higher quality of the firms funded after Brown's election means gives them an advantage over the firms funded directly before. This leads to Hypothesis 3:

Hypothesis 3: Because of changing expectations, (a) VCs should be less willing to fund clean energy startups after Brown's election and (b) the clean energy startups that received Series A funding in the months directly prior to Senator Brown's election should be less successful than the clean energy funded just after Senator Brown's election.

To test Hypothesis 3.a, we look at the share of VCs' portfolio going to clean energy startups. To test Hypothesis 3.b, we implement a difference-in-differences estimation using Equation 2.2:

$$Success_i = \beta_0 + \beta_1 CleanEnergy_i + \beta_2 PreBrown_i + \delta CleanEnergy_i * PreBrown_i + \eta_{state} + \epsilon_i \quad (2.2)$$

Because we only consider firms that receive Series A funding at similar points in time, we are comparing firms at similar levels of development. Thus, we can consider multiple measures of success. *Success* can be the probability of securing follow-on Series B or C funding, an exit dummy, a cash-on-cash measure or the value at exit. We compare the effect the shift in expectations had on *Clean Energy* startups and ICT startups. The latter industry is a good control group due to its many observations. Moreover, ICT startups should not be affected by changes in climate policy expectations while still controlling for changes in the broader economic outlook. We only include in our *Clean Energy* group explicitly clean energy startups rather than on our broader clean energy & EVs category to focus on the startups most affected by the tentative climate bill. We argue that the climate policy change would have had more impact on the startups directly involved in producing and storing clean energy and less on EVs and startups in energy efficiency or grid management startups (see Table 2.2). We include all startups explicitly engaged in clean energy, storage, renewable, solar or wind. We also include all startups with a "cleantech" and "energy" tag. Our sample includes 34 explicitly clean energy startups having secured their first Series A funding in the 9 months before and after Brown's election. In addition, and because we want to maximize our sample size, we manually verified the activity of 15 startups in either energy efficiency and grid management or non-energy cleantech to include those that might fit into clean energy despite the lack of an explicit tag. This allows us to include one additional firm developing a renewable and CO2 absorbing alternative to petrochemicals, as the value of the emissions-reducing benefits of this technology are minimal without carbon emission regulations.

We run regressions using either a 6 or 9-month window around Brown's election. The 6-month window is more precise but has less observations than the 9-month window. The period between the 1st and 14th of January is not included in either the *PreBrown* or post-Brown period. Polls conducted at the beginning of January were already hinting at the potential victory of Brown and thus investors' outlook might already have shifted. We choose the 14th of January as the threshold because the seven polls after the 13th of January have Brown winning by a healthy margin (3% to 15% – except the smallest poll that has them tied).[16] We assume that our treatment effect, δ, only captures the effect of the change in expected demand for clean technologies as ICT startups control for potential changes to the overall business outlook and are not directly affected by Brown's election during the timeframe of our experiment. Unlike the previous section, we do not include biotech firms in the regression, as the potential profitability of biotech firms may have been affected by changing expectations for the success of the Affordable Care Act, which had not yet passed the Senate and eventually was passed using reconciliation to avoid a potential filibuster.

Before estimating δ, the *treatment* effect of Brown's election on startups' success, we first look at the prevalence of clean energy startups within VC portfolios in the 9 months before and after Brown's election. We find that 4.9% of the startups that received their first Series A funding during the pre-Brown period were clean energy, compared to only 3.7% in the post-Brown period. This 24.5% drop suggests that VCs did indeed reduce their investments in clean energy in the 9 months after Brown's election.

Turning to the impact of this shift in expected demand on startup quality, Table 2.5 implements Equation 2.2 for the 6-month window in Panel A and Panel B shows that the results hold using the 9-month window. In Column (1) we can see that startups that received their first Series A just before Brown's election had a significantly lower probability of securing follow-on Series B and C than those funded directly after. Columns (2) and (3) show that the clean energy startups funded just before the unexpected election of Brown has a slightly lower chance of getting an IPO, but this effect is not statistically significant. Columns (4) to (7) use more refined measures of success that better capture startups' outcome; the logarithm of the CoC and the discounted CoC, whether the startup returned at least 5 times the invested capital as well as the logarithm of the exit value. All these success variables show the same thing: startups funded pre-Brown were less successful than those funded post-Brown.

[16]We use the 14th rather than the 19th (i.e., the day of the election) as the outlook started to change before the election. It also allows us to include a few more startups in our sample. The results are robust to using both the 14th or 19th as the threshold.

Table 2.5: Difference-in-differences estimation, the effect of Brown's election

	(1) Series B/C	(2) Exit	(3) IPO	(4) CoC (ln)	(5) disc. CoC (ln)	(6) 5x returns	(7) Exit val. (ln)
Panel A: 6-month window							
Clean energy	0.188	-0.344***	0.00271	-0.465***	-0.371***	-0.0258	-4.343***
	(0.132)	(0.125)	(0.0530)	(0.153)	(0.125)	(0.0477)	(1.142)
Pre-Brown's election	0.0360	-0.0286	0.0108	0.283	0.231	0.0465	2.026
	(0.0556)	(0.0554)	(0.0291)	(0.185)	(0.151)	(0.0353)	(1.290)
Clean energy pre-Brown	-0.415**	-0.0530	-0.0762	-0.450*	-0.354*	-0.131**	-3.897**
	(0.203)	(0.183)	(0.0603)	(0.240)	(0.194)	(0.0638)	(1.865)
Panel B: 9-month window							
Clean energy	0.120	-0.322***	0.0173	-0.240**	-0.179*	0.0413	-2.673**
	(0.115)	(0.0955)	(0.0471)	(0.114)	(0.0954)	(0.0522)	(1.068)
Pre-Brown's election	0.0103	0.00861	0.0314	0.252*	0.193*	0.0472*	2.168**
	(0.0441)	(0.0443)	(0.0215)	(0.135)	(0.110)	(0.0267)	(0.980)
Clean energy pre-Brown	-0.252	0.163	-0.0914*	-0.445**	-0.363**	-0.153**	-3.325
	(0.171)	(0.160)	(0.0509)	(0.187)	(0.152)	(0.0596)	(2.312)
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations Panel A	379	379	379	238	237	353	240
Observations Panel B	569	569	569	353	352	518	358

Note: This table presents results of OLS regressions using difference in differences estimations (DiD), using Equation 2.2. The sample includes *explicitly* clean energy and ICT startups that received Series A funding during our 6- or. 9-month windows. In panel A, startups are *pre-Brown* (i.e., *Pre-Brown's election = 1*) if they received their first Series A from the 30th of June 2009 to the 1st of January 2010 and post-Brown if they received Series A from the 14th of January to the 15th of July 2010. In panel B, the pre- and post-Brown time windows are from the 31th of March 2009 to the 1st of January 2010 and from the 14th of January to the 15th of October 2010, respectively. *Clean energy pre-Brown* captures our treatment effect. In Column (1), the dependent variable is whether startup *i* secured Series B or C, in (2) whether it exited, in (3) whether it went public, in (4) the logarithm of its cash-on-cash return, in (5) the discounted value of the logarithm of its cash-on-cash return, in (6) whether it returned at least 5x to its Series A investors and in (7) the logarithm of its exit value. The level of observation is at the firm-level. Robust standard errors are in parentheses. Asterisks denote * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

One important caveat with the results in Columns (4) and (5) and (7) is that they depend on exit values that are not often disclosed. Indeed, 11 out of the 35 clean energy startups funded during the 18 months of our event study managed to exit. However, we only have the actual exit value for the one startup that did an IPO and 2 out of the 10 companies that got acquired. To address this problem, we include Columns (1) and (6), as they are dummy variables available for all startups.

Overall our results show that demand-side policies matter for clean technologies. We show that after a negative demand shock, VCs allocate a lower share of their portfolio to clean energy companies. In addition, we find that the startups that do get funded tend to outperform those that got funded just before the negative demand shock, when funding requirements were lower. These results are evidence that when VCs foresee lower policy support, they lower their expectations of returns and, as a result, become less willing to fund clean energy startups and increase the quality threshold they must meet. We, therefore, argue that politicians wishing to support clean energy startups must first create a policy environment in which there will be dynamic demand for their products.

2.4.3 The role of the low potential for outsized returns

On top of lackluster demand, we document another barrier to VC success in clean energy: the inability to earn high margins because of a lack of product differentiation (Nanda et al., 2015). VCs like to invest in sectors, like ICT or biotech, where they can expect their successful bets to yield phenomenal returns that will compensate for the risk of funding many early-stage ventures. These great returns stem partly from the fact that ICT and biotech startups can build moats around their product, which grant them significant market power. In ICT, network effects and high switching costs mean that once a company has established its dominance, it is very difficult for competitors to challenge it (Katz and Shapiro, 1992; Gallaugher and Wang, 2002). In biotech, patents protect new drugs and other products. In clean energy, companies have to accept the thin margins of competitive energy markets. Additionally, renewable energy companies cannot differentiate the electricity they produce from the electricity produced using coal or gas. Overall, the lower market power of clean energy firms means that they cannot reach the profitability observed in ICT and biotech. In contrast, EVs' ability to differentiate their product – think Tesla – gives them more market power than energy producers.

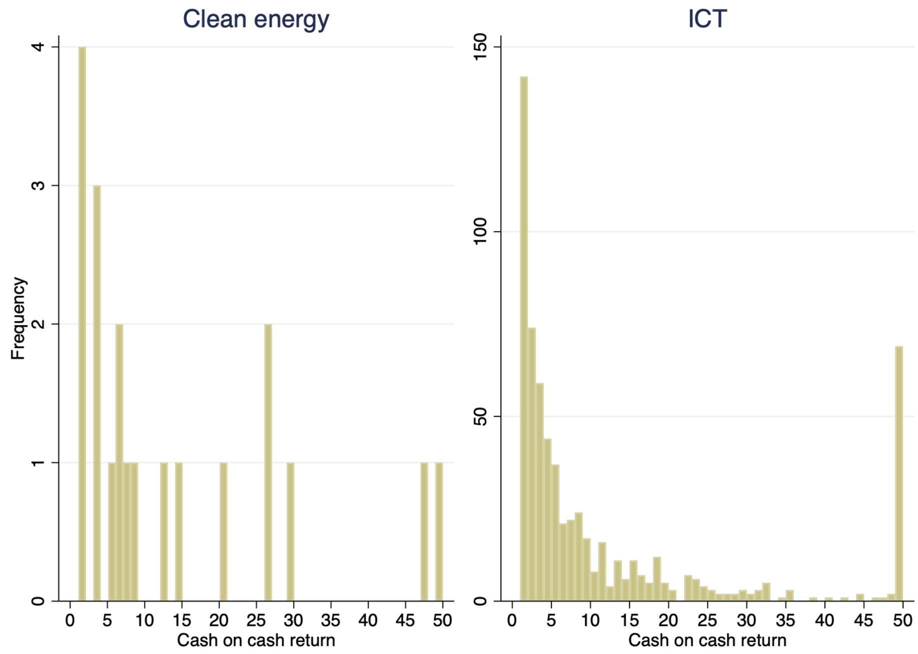
In this section, we look at the returns of Series A investors in our different industries to see whether the data supports the argument that the energy sector does not offer outsized returns. To answer this question, we look at the Cash-on-Cash (CoC) return and exit value of startups with Series A funding. First, we find that clean energy startups are less likely to provide massive returns to their investors. Indeed, 2.1% of ICT startups and 2.6% of biotech startups that receive Series A return at least 10x to their Series A investors (i.e., have a CoC above 10). This number stands at only 1.6% for clean energy startups.[17] The difference, especially between clean energy and ICT, is even more striking when looking at outsized returns. Figure 2.7 shows the distribution of returns in clean energy and ICT for the startups that have at least returned the invested capital to Series A investors. We group all companies that have returned more than 50x in the rightmost bin. In clean energy, the biggest success, Opower,[18] returned 56x the invested capital to its Series A investors. At the same time, 36 digital and 3 biotech startups have returned more than 100x the invested capital. Based on past returns, it is fair to say that early-stage investors in clean energy cannot expect to hit the home run successes that Snowflake (704x), Uber (781x), Facebook (832x) or Roblox (4651x) have been.

More formally, we estimate the probability that energy & EV startups will return at least 5x or 10x the initial Series A investment (relative to ICT startups) or that it recorded an exit value

[17] We do not show the EV share as observations are rather low. According to the available data 1 out of 39 EVs returned more than 10x (2.6%), but exit values were undisclosed for 4 companies.

[18] Opower is a software company that uses AI and behavioral science to optimize its customers' energy consumption. Its ICT features and self-reinforcing use of data, which grants them market power, could explain why it has proven so much more profitable to VC investors that the typical clean energy company.

Figure 2.7: Returns for Series A investors in clean energy & ICT



Note: This figure shows the cash on cash return of startups with Series A funding that returned at least 1x, by sectors. Startups that return above 50x are grouped in the 50x bin to limit the size of the x-axis

above $1B or $5B.[19] Table 2.6 shows that clean energy startups with Series A funding have consistently lower probabilities of being big successes than ICT.[20] The biotech firms have a slightly higher probability of being successful (returning at least 5x) than ICT but a lower probability of being a home-run success (above $5B exit).

[19]These exit values are relatively rare. Only 28.2% and 6.5% of companies that IPO reach a valuation of $1B and $5B respectively. For acquisitions, which represent two thirds of exit events, these numbers stand at 10.1% and 0.9% respectively. Overall, 35.4% of companies that exit return at least 5x and 21.4% return at least 10x.

[20]We do not discuss the EV results are they are based on only three startups with at least 5x.

Table 2.6: Probability of outsized returns across sectors

	All Series A startups			
	(1)	(2)	(3)	(4)
	5x returns	10x returns	Exit above 1B	Exit above 5B
Energy	-0.0274**	-0.0186**	-0.0198***	-0.0047***
	(0.0113)	(0.0089)	(0.0041)	(0.0011)
Energy digital	-0.0183	-0.0212**	-0.0205***	-0.0041***
	(0.0175)	(0.0105)	(0.0024)	(0.0009)
Electric vehicles	0.0312	0.0469	0.0476	-0.0052***
	(0.0543)	(0.0553)	(0.0498)	(0.0014)
Biotech	0.0279***	-0.0006	-0.0050	-0.0032***
	(0.0083)	(0.0057)	(0.0042)	(0.0011)
Constant	0.0345	0.0101	-0.0230***	-0.0010
	(0.0249)	(0.0176)	(0.0083)	(0.0015)
Series A year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	7759	7759	8698	8698
R^2	0.026	0.016	0.017	0.005

Note: This table presents results of OLS regressions. The sample includes clean energy & EVs, ICT and biotech startups that received Series A from January 2003 to December 2016. In Column (1), the dependent variable is whether the startup returned at least 5x the paid-in capital to its Series A investors and at least 10x in Column (2). In Column (3), the dependent variable is whether the startup exited at a valuation of at least $1B and at least $5B in Column (4). The level of observation is at the firm-level. Robust standard errors are in parentheses. Asterisks denote * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Overall, our results suggest that lackluster demand is a key factor behind the failure of VC in clean energy since the boom of 2005-2009. As we do not observe a world in which externalities have been fully addressed, we cannot definitively state that oversized returns would not be possible in clean energy if demand was addressed. However, the inability for many clean energy firms to product differentiate makes it likely that, for at least some clean energy products, addressing demand may not be sufficient to make VC successful for clean energy. On the other hand, we do not find evidence that capital-intensity and long-time horizon are a key factor behind the disappointing performance of VC in clean energy.

2.5 The role of governments in supporting clean technologies

Our results suggest that a reliable way through which governments could support early-stage investments and innovation in clean energy technologies would be to stimulate demand to increase startups' profitability. A carbon tax or cap-and-trade system would create higher demand for clean energy & EVs and make startups in these sectors a more attractive acquisition target. Both of which would increase VCs' willingness to invest in early-stage clean energy & EVs startups. However, while there is a consensus among economists that governments should use market-based instruments to internalize environmental externalities and incentivize investments in clean technologies, actually implementing them is an arduous political task. Moreover, boosting clean energy firms' profits by giving them more market power would not be optimal, as competition is also an important driver of innovation.

In an ideal world, governments would therefore support private investments by stimulating demand. Until this happens, investing in early-stage clean energy will prove challenging. Nevertheless, there is still a great need for more investments in early-stage clean energy startups and providing public funding is easier politically than implementing demand-side policies such as a cap-and-trade system. We, therefore, argue that it is worth looking into the role and performances of public investors.

Public investors should not be seen as a substitute to VC and its unmatched financial capabilities. According to the National Venture Capital Association, investors poured $156.2B into U.S. startups in 2020. Instead, public investors should be seen as a complement to VC. Tellingly, the success of public programs is often measured by whether recipients of early public awards are able to attract subsequent VC funding (e.g., Howell, 2017).

As a complement to VC, public investors can provide funding either before or after Series A rounds. Table 2.7 shows that the amount offered varies greatly depending on the timing of the public involvement. Indeed, while public investors give $0.72M on average to clean energy startups pre-Series A, they provide $15.4M post-Series A, a more than twentyfold increase. These numbers highlight that public investors play a very different role in the early and late stages of a startup's lifecycle. Early-stage funding is designed to help clean energy startups prove the potential of their technology and attract VC money, thus bridging the early technological valley of death. The much more generous later-stage funding can help VC-funded energy startups scale up and survive the commercialization valley of death.

Table 2.7: Average size of public investments pre- and post-Series A

Clean energy	in 2020 million USD	Public Amount	Series A Amount
n = 23	Public before Series A	0.72	6.1
n = 32	Series A before Public	15.4	10.4
Digital clean energy	in 2020 million USD	Public Amount	Series A Amount
n = 11	Public before Series A	0.7	4.6
n = 11	Series A before Public	12	5.2

Note: This table displays the average size of public and Series A investments for startups that received public money before their first Series A round and vice versa, by clean energy category. Two outliers that had an outsized impact on the average were removed. A green infrastructure investment firm that received $1B from a public pension fund after its Series A and a firm that received $100M in a pre-Series A public investment.

In this section, we test the effectiveness of both early and late-stage public investments in helping their recipient bridge their relevant valley of death. To do so, we compare the performances of clean energy startups that receive public funding to those that receive similar, but private, funding.

Early-stage public involvement as a way to bridge the first valley of death

While early-public money is limited in size, as shown in Table 2.7, it can nonetheless prove helpful. Cheap, early-stage public investments can indeed signal value to private investors and provide the resources necessary to prove the viability of their technology to Series A investors. To test that, we estimate whether receiving early-stage public money in the first two years following the startup's launch is associated with better outcomes both in the short (i.e., securing Series A) and the long run (i.e., exit outcomes). However, we need to control for the fact that receiving early money is already a sign of startup quality. We therefore include receiving private seed funding in our regression. We argue that seed funding is a good benchmark for early public money. First, the amounts provided are very similar. The average size of an early public funding (within two years of launch) for a clean energy company is $1.3M versus $1.4M for early seed money.[21] Second, like early-stage public support, seed funding typically represents the first official outside capital that a startup raises and also helps startups attract Series A investors.[22]

Table 2.8 looks at clean energy & EVs as a whole. *Early public money* and *Early seed money* are equal to 1 if a startup received either source of funding during its first two years of existence. In Column (1), we see that both early public money and early seed are associated with a higher chance of securing VC funding. The effect of public money is slightly bigger; early public money is associated with an increase of 21.4 percentage point in the probability of securing Series A versus 11.4 percentage points for early seed. Columns (2), (3) and (4) show that conditional on having received Series A, having both public money and seed are not correlated with late-stage success. Our results suggest that early public money is helpful to bridge the first technological valley of death and secure Series A, maybe even more so than seed funding. However, we see that once they secure Series A, the early public involvement has no bearing on late-stage success. In other words, it does not help clean energy startups bridge the second, commercialization, valley of death. However, this is not the direct aim of early-public money, unlike later-stage public support.

[21] Note that this number slightly differs from the numbers in Table 2.7 because the sample is not identical ("within two years of launch" versus "before series A"). Additionally, one clean energy firm received $60M in public money within its first two years and is removed from the average as the 77 other clean energy firms only received a bit more than $100M collectively.

[22] See Crunchbase's own website for a definition of most types of funding: https://support.crunchbase.com/hc/en-us/articles/115010458467-Glossary-of-Funding-Types

Table 2.8: The relationship between early public money and subsequent success

	Funding	Success conditional on Series A		
	(1)	(2)	(3)	(4)
	Series A	Exit	IPO	Acquired
Early public money	0.2125***	-0.0201	0.0109	-0.0369
	(0.0503)	(0.0699)	(0.0348)	(0.0609)
Early seed money	0.1063***	0.0119	-0.0202*	0.0296
	(0.0239)	(0.0433)	(0.0107)	(0.0429)
Constant	0.2904	0.2801*	0.1355	0.1959
	(0.2755)	(0.1443)	(0.1101)	(0.1244)
Launch year FE	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes
Observations	4160	510	510	510
R^2	0.071	0.189	0.086	0.185

Note: This table presents results of OLS regressions. In Column (1), the sample includes clean energy & EVs that launched from January 2003 to December 2016 and the dependent variable is whether the startup received Series A. In Columns (2), (3) and (4) the sample only includes startups that received Series A funding and the dependent variables are whether the startups exited, got acquired or went public. *Early public money* and *Early seed money* are equal to 1 if startup *i* received either source of funding in the first two years after its launch. The level of observation is at the firm-level. Robust standard errors are in parentheses. Asterisks denote * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Late-stage public involvement as a way to bridge the second valley of death

We, therefore now turn to look at the performance of public actors when supporting post-Series A clean energy startups trying to scale up their business. In particular, we are interested in estimating how public investors compare to private investors. This is important as, as shown in Table 2.7, public investors provide sizeable amounts to clean energy startups after they secure a Series A round.

As always, estimating the effect of public funding is difficult given the lack of a counterfactual. At its most effective, the public funding goes to clean energy startups that would not have secured late-stage private funding otherwise. In Table 9, Columns (1), (2) and (3) show that companies that receive series A and then public money have a significantly higher probability of exit than those that receive Series A and then nothing. The results suggest that, at best, if private funding is unavailable, public money can help. At its least effective, public funding goes to startups would have been funded by VCs anyway. Columns (4), (5) and (6) show that companies that receive public money after Series A have similar exit probabilities than those that receive VC funding after Series A. At worst, public funding does no better but also no worse than receiving late-stage venture capital money. Overall, companies that receive private and public support fare at least as well as those that only rely on private investors.

However, these results must be viewed in context with the results from the previous section, which showed that venture capital investments in clean energy did much worse than in other sectors. Public funding cannot address alone the overall poor performance of investors in clean energy. As such, public investments are unlikely to bridge the commercialization valley of death for clean energy startups by themselves. However, if policies are put in place to address demand-side problems for clean energy, it seems likely that public sector money will have some role to play. The lower bound estimates above assume that none of the late public sector investments are additional – all of these companies would eventually have received private late round VC funding. But even if demand-side issues are solved, many clean energy firms will still have difficulty differentiating their products in a way that makes the outsized returns demanded by VC investors possible. Thus, it is likely the case that a role for public investment will remain.

Table 2.9: The impact of late-stage funding on exit probabilities

	Upper bound effect			Lower bound effect		
	(1) Exit	(2) IPO	(3) Acquired	(4) Exit	(5) IPO	(6) Acquired
Post-Series A: Public vs Nothing	0.1863^{*}	0.0880^{*}	0.1228			
	(0.1079)	(0.0507)	(0.1022)			
Post-Series A: Public vs VC				-0.0334	0.0731	-0.0861
				(0.0924)	(0.0557)	(0.0921)
Constant	0.3267	0.2389	0.0818	1.3282^{***}	0.0829	1.1803^{***}
	(0.2521)	(0.2429)	(0.1571)	(0.2893)	(0.1683)	(0.2900)
Launch year FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	149	149	149	202	202	202
R^2	0.277	0.277	0.238	0.142	0.208	0.145

Note: This table presents results of OLS regressions. In Columns (1), (2) and (3), we compare the exit performances of clean energy & EVs startups that received Series A and then no more funding to the exit performances of clean energy & EVs startups that received Series A and then public funding. In Columns (4), (5) and (6), we compare the exit performances of clean energy & EVs startups that received Series A and then VC funding to the exit performances of clean energy & EVs startups that received Series A and then public funding. The sample includes startups that received Series A from January 2003 to December 2016. The level of observation is at the firm-level. Robust standard errors are in parentheses. Asterisks denote * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.6 Conclusion

Using data from 250,000 companies launched in the U.S. between 2000 and 2020, we evaluate the performance of venture capital investments in clean energy. Most importantly, our results show that demand matters. VC investors adjust their behavior in response to changes in expected future demand. The downward path of policy support can therefore explain, at least in part, the post-2009 trajectory of VC in cleantech. The lowered expectation of future policy support, because it reduced VC's willingness to fund clean energy ventures, contributed to the end of the cleantech boom.

We also highlight a second reason for the failure of cleantech VC. Unlike ICT ventures, clean energy startups are less likely to become home run successes. The lack of network effects, the lesser reliance on patents and lower product differentiation make it harder to keep competitors at bay and earn high margins in clean energy. This prevents the formation of winner takes all markets, which makes betting on ICT startups so attractive. While stimulating demand will increase average returns, it will not alter clean energy's market structure. Thus, venture capital is not a one-sized-fits-all solution. Even after demand-side issues are solved, clean energy products are more likely to be attractive to VC investors in cases where product differentiation is possible. A good example is Nest, a smart thermostat that optimizes energy consumption. Nest Labs managed to differentiate its product and attain high levels of brand recognition, and, as a result, got acquired by Google for $3.2B in 2014.

Finally, while the capital intensity and the long development timeframe of energy startups are often cited as barriers to raising funds for energy firms, we argue that they are not the main factor behind the failure of VCs in clean energy. Indeed, after having fared better in the early stages of the cleantech bust, non-capital-intensive digital energy startups eventually also experienced a fall in their profitability and in their ability to secure VC funding. At the same time, biotech and EVs, another cleantech sector, have been able to attract investors in recent years despite being capital-intensive. Moreover, if these startups characteristics were indeed to blame for the limited funding going to clean energy, the few, high-quality, startups that do get funded should yield high returns. However, this is not what we find: there is no evidence that VC investors in clean energy are "leaving money on the table".

In this context, the role of governments is clear: implementing demand-side policies to support clean energy investments and innovation. Direct financial support to clean energy startups is indeed less likely to have a major impact since liquidity constraints do not seem to be the main reason behind the lack of success of clean energy investments. However, Democrats failed to pass a comprehensive climate bill in 2009-2010 when they controlled both Congress and the Presidency. This failure highlights the political difficulties in implementing climate policies. On the other hand, providing financial support to startups remains a popular policy tool. We show that public money can complement VC money in two ways. In the early stage, public investors provide small grants that help startups bridge the technological valley of death and secure Series A funding. In later stages, public investors provide much larger amounts that help energy startups expand and scale their business but does not fully address the commercialization valley of death. More importantly, we show that, as a complement to VC, public investors have not fared worse than their private-sector counterparts. Our results are in line with previous findings that public support is either helpful (Howell, 2017) or, to the very least, not worse than other sources (Goldstein et al., 2020). However, this is setting a low bar given the poor performance of VCs in clean energy. Without a positive demand shock, neither public nor private investors are likely to be consistently successful when funding clean energy startups. Only by creating the conditions for robust demand for green products will governments be able to address the valley of death problem. After addressing demand issues, public sector funding for startups will be most successful if it targets firms developing products where product differentiation, and thus outsized returns, is most difficult.

The role of demand in the second cleantech boom

The importance of strong demand can help us understand the second boom in cleantech VC and provide insights as to which investments are more likely to succeed. First, government policies are becoming more supportive than they were in 2009. In November 2021, President Biden signed the Infrastructure Investment and Jobs Act that contains more than $70M in support of clean energy and better electrification of the United States. This first bill might be followed by a second, larger, spending plan to fight climate change. Regardless of the outcome of this second bill, government support is strengthening. For instance, Biden signed an executive order declaring that half of the new cars sold in 2030 should be EVs. Finally, the COP26 in Glasgow saw the United States and many other countries committing to a net-zero emissions economy by no later than 2050. This more supportive policy environment is boosting demand for all clean technologies.

However, VC investments in clean energy are likely to continue lagging behind other cleantech sectors like EVs (see Figure 2.1). Indeed, VCs have learned to focus on cleantech sectors with already high demand, notably these sectors in which using the "clean" technology requires little sacrifice (TheEconomist, 2021). Some sectors like EVs or sustainable food (e.g., plant-based meat) have few barriers to mass adoption, helping fueling demand. By contrast, regardless of its low cost, renewable energy is dependent on other energy sources because of its intermittency. As long as electricity cannot be stored on a massive scale or easily traded across regions, demand for clean energy will suffer and this even with a price on carbon. Finally, sectors like EVs and sustainable food can differentiate their products, which improves their profitability. The ability of firms like Tesla to manage their brand and bolster demand for their particular EVs allows them to generate markups unattainable in the more competitive renewable energy industry. These attractive features have translated into several recent home run successes in cleantech. Tesla returned 20x the invested capital during its IPO in 2010, but its valuation has since increased by 647x to a market value of around $1.1T. This has whetted the appetite for investment in the area. Nikola Motor Company, a producer of electric trucks, returned 125x the paid-in capital to its Series A investors when it went public in 2020. A similar pattern has been visible in sustainable food. Impossible Foods, a maker of plant-based meats is now valued at $7B and should go public in 2022, returning 82x to its early-stage investors. These companies have proven that some cleantech sectors are able to produce outsized returns, fueling this new cleantech boom.

Examples like these show that investors are increasingly able to identify cleantech firms for which venture capital can succeed. Thanks to a policy environment expected to become increasingly supportive in the future and a focus on sectors with both relatively higher demand and brand loyalty, the second boom in VC may result in better outcomes than its predecessor. However, the smaller overall size of the second boom is indicative that clean energy startups are likely to continue suffering from a funding gap. Their lack of product differentiation and the absence of carbon pricing will continue to make them relatively unattractive to profit-driven investors. As more innovation in clean energy is needed to address climate change, governments should do more. Rather than investing themselves in startups bound to struggle through the valleys of death, governments should first implement demand-side policies that make investing in clean energy more viable. Only then should they turn to financing new firms. That support should focus on the firms in cleantech markets where product differentiation is difficult, as such firms are least likely to attract private sector support.

3 Heard the News? Environmental Policy and Clean Investments

This chapter is written in collaboration with Joëlle Noailly and Laura Nowzohour.[1]

Abstract:

We build a novel news-based index of US environmental policy and examine how it relates to clean investments. Extracting text from ten leading US newspapers over the last four decades, we use text-mining techniques to develop a granular index measuring the salience of US environmental policy (EnvP) over the 1981-2019 period. We also develop a set of additional measures, namely an index of sentiment on environmental policy, as well as various topic-specific indices. We validate our index by showing that it correctly captures trends and peaks in the evolution of US environmental policy and that it has a meaningful association with clean investments, in line with environmental regulations supporting growing opportunities for clean markets. In firm-level estimations, we find that the salience of environmental policy in newspapers is associated with a greater probability of receiving venture capital (VC) funding for cleantech startups and reduced stock returns for high-emissions firms most exposed to environmental regulations. At the aggregate level, we find in VAR models that a shock in our news-based index of renewable energy policy is associated with an increase in the number of clean energy VC deals and in the assets under management of the benchmark clean energy exchange-traded fund. Overall, our EnvP index provides substantial information on environmental policy and can help assist the policy and financial community in understanding how these regulations are perceived by investors — providing many avenues for future research.

[1] The three co-authors contributed equally to the core of the paper: the research question, framing, writing and labelling of newspaper articles. Hired research assistants also contributed to the labelling. The doctoral candidate and Laura Nowzohour (L.N) implemented the SVM and LDA (machine learning). Joelle Noailly (J.N) is the main contributor to the introduction. L.N is the main contributor to the stock market analysis (in Section 3.4) and the sentiment analysis (Section 3.3.1). The doctoral candidate is the main contributor to the venture capital analysis (in Section 3.4) and the evaluation of the index (Section 3.2.2)

3.1 Introduction

Stabilizing global warming and achieving net zero emission targets requires implementing a wide array of climate and environmental regulations. Such policies are essential to shift incentives of economic actors towards investments in clean products, technologies or firms. Yet, the salience of environmental policy on the political agenda tends to fluctuate over time. The US election of President Joe Biden in November 2020 illustrates how climate policy can gain renewed prominence over the course of a few months. We contribute to the literature aiming to measure the effectiveness of environmental policy in fostering clean investments by developing a novel news-based index of environmental policy relying on textual analysis of newspapers. We apply text-mining techniques to articles from ten leading US newspapers to construct a general index measuring the salience of US environmental and climate policy, available on a monthly basis over the 1981-2019 period. Furthermore, we develop a set of additional measures, namely an index of general sentiment on US environmental policy and 25 topic-specific indices, allowing us to distinguish between various policy issues such as international climate negotiations or renewable energy policy, among other topics.

We evaluate our general environmental policy (EnvP) index by showing that it captures significant policy events in the history of US environmental regulations and that it co-moves with the evolution of environmental policy stringency. We further validate the index by showing that it has a meaningful association with financial investments most exposed to environmental regulations. More specifically, we find in firm-level estimations that our EnvP index is associated with a greater probability of cleantech startups receiving VC funding and reduced stock returns for high-emissions firms most exposed to environmental policy. Moreover, we find in VAR models that a shock in our news-based index is associated with an increase in the number of clean energy deals at the macro level and an increase in the assets under management of the main clean energy exchange-traded fund. These results suggest that increased environmental policy salience fosters growing opportunities for clean markets.

Our news-based measure of environmental policy provides several complementary insights to existing quantitative indicators of environmental regulations. First, in contrast with many studies assuming that consumers and investors have full knowledge on environmental policy instruments, our index directly builds on the volume of (relatively low-cost and easily accessible) information contained in news which is likely to affect investors' perceptions and decision-making.[2] Hence, our analysis postulates that the salience of environmental policy matters for the incentives of economic actors.[3]

Second, because news arrive daily and are available over long periods of time, our news-based

[2]While we could argue that professional investors may rely on more sophisticated information channels, such as business news or social media, newspapers present the advantage of being available for much longer time periods and to be relatively low-cost compared to other media. In addition, we consider newspapers, such as *The Wall Street Journal*, which target the audience of investors.

[3]Huse and Koptyug (2021) show for instance that when consumers do not pay attention to less salient taxes, they are more likely to make an incorrect valuation of vehicle costs.

index is a significant improvement over existing indicators of environmental policy computed on an annual basis. Our approach provides a continuous tracking of environmental policy over time at a high frequency (monthly and quarterly time series), making it possible to measure immediate market reactions and to better address unobserved heterogeneity in empirical work (Ghanem and Smith, 2020). By contrast to event studies, our index is able to capture long-lasting dynamics of the policy process (e.g. announcements, delays, revisions) and their potential economic impacts on markets. Extracting information from newspapers may also provide additional information on the regulatory context – such as details on implementation, controversies and sentiment – which are not typically captured by standard indicators.

Third, our index is available at various levels of aggregation (generic or topic-specific) helping to address the challenge of the multidimensionality of environmental policy. Regulations are often introduced as a 'package' of policies covering multifaceted aspects (such as the Green New Deal) and governments typically regulate the polluting activities of households and industries across many sectors on a wide range of pollutants. Summarizing and quantifying these various aspects into meaningful composite indicators is a very challenging task in empirical work (Brunel and Levinson, 2016). By using machine learning techniques on the rich amount of text provided by news articles, we are able to disentangle (latent) information on various sub-clusters of environmental policy issues and to build topic-specific indices, tracking for instance the salience of renewable energy policy or international climate negotiations over time.

Our approach raises several concerns about significance, accuracy and potential bias which we evaluate in several ways. A first concern relates to what our index actually measures and how it relates to policy stringency. As environmental policy becomes more stringent, economic agents will have greater (lower) incentives to invest into clean (dirty) markets. In a similar fashion, an implicit assumption in our work is that an increase in the volume of environmental policy news raises the awareness of economic agents about existing regulations and growing opportunities in clean markets, leading them to increase their clean investments.[4] We first verify that the salience of environmental policy in newspapers correlates with the evolution of regulatory stringency over time. We further validate our news-based index by showing that it is associated with financial investments – as proxied by venture capital investments and stock returns – in a manner that is consistent with environmental regulations opening up growing opportunities in clean markets. An immediate concern when focusing on the relative volume of news is that our index may inaccurately capture negative discussions about environmental policy – e.g. relating to the high costs of environmental regulations leading to opposition by lobbyists or regulatory rollbacks – giving rise to perceptions of a decline in stringency. We deal with this by showing that our results remain robust when adding controls for a measure of environmental policy sentiment.

Additional concerns relate to the accuracy of our index in capturing the salience of environ-

[4]Note that we do not attempt here to separate the specific impact of regulation from the impact of media coverage – a highly challenging task. Instead, both are entangled into our index.

mental regulations. Environmental problems (and their semantics) evolve over time and we may be worried about missing out important policies. We address this issue by relying on machine learning approaches, rather than manual labelling or refined keyword searches. These techniques present the advantage of easing the processing of large amounts of text and of uncovering latent patterns without imposing too much structure on the text – although we invariably provide a minimum level of critical data to train the algorithms and inform the models. Machine learning approaches also make it possible to quantify measurement errors and to assess the performance of various algorithms. We find that our supervised learning model predicts environmental policy articles and captures relative trends over time relatively well.[5]

An additional important concern is that our index may be affected by media bias. The production of news is the result of an equilibrium between demand- and supply-side factors. On the demand side, newspapers feed news according to the preferences of their readers towards specific topics (Gentzkow and Shapiro, 2010; Mullainathan and Shleifer, 2005). As readers become more aware about environmental problems, they may demand more reporting about these and the policies to address them. On the supply side, economic incentives, competitive pressures between topics, journalists' norms[6], and communication from policy organizations[7] all affect which topics are being covered on a daily basis (Baron, 2006). Competition between media outlets and readers' heterogeneity, however, tend to provide incentives to increase accuracy in the reporting of news and contribute to mitigate media bias (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006). Accordingly, we include a wide range of newspapers in our analysis and investigate potential bias due to partisan readerships. As macroeconomic and political factors also influence the reporting of environmental policy news – for instance public support and interest for environmental policy typically falls during recessions (Kahn and Kotchen, 2011), we control for business and political cycles via time fixed effects in our empirical estimations.

Finally, our validation exercise examining whether our news-based index has a meaningful impact on clean investments may be biased by potential endogeneity issues. Journalists may be reporting relatively more about environmental policy when clean markets are growing, for instance under the influence of green lobby groups. Omitted variables such as technological change may also affect both media coverage and market outcomes. As such, establishing causality in the absence of natural experiments is highly challenging. To make progress, nonetheless, our identification strategy relies on a differentiation of firms and sectors accord-

[5]Due to inevitable measurement errors, our index does not pretend to retrieve the whole universe of environmental policy articles. Trends are however correctly identified as long as the distribution of environmental policy news remain constant over time.

[6]There is for instance evidence that in the US journalists' ideological preferences and norms gave a misrepresentation of the scientific consensus on climate change (Boykoff and Boykoff, 2007; Brüggemann and Engesser, 2017; Shapiro, 2016), which could slant reporting about environmental policy in a negative way.

[7]Muehlenbachs et al. (2011) find that controversial press releases from the Environmental Protection Agency on enforcement actions or regulatory changes were issued more often on Fridays, a time when news has the least impact on media coverage.

ing to their exposure to environmental policy. Akin to a difference-in-difference approach, we estimate whether our news-based index has a differentiated impact on investments in venture capital or stock returns of firms most exposed to environmental policy, as defined by sector of activity or emission levels. These (micro) firm-level estimations help us build suggestive evidence towards causal inference. We further test the robustness of our results to many fixed effects specifications in order to reduce the influence of confounding factors. At last, we examine the dynamic relationship between our index and aggregate clean investments in (macro) VAR settings. Although establishing causality is arduous, the persistence of our results across various measures of clean investments, a large range of robustness specifications, and at both micro and macro levels, can be interpreted as indicative that environmental policy – as measured by our news-based index – helps promote the growth of clean markets.

Our study contributes to the existing literature in several ways. First, our approach contributes to a broad literature discussing how to construct meaningful indicators of environmental policy (Brunel and Levinson, 2016). Previous literature has solved the multidimensionality issue of environmental policy either by using a specific regulation as an indicator of overall environmental regulation – such as grams of lead-content per gallon of gasoline as in Cole and Fredriksson (2009) – or by constructing composite indicators based on a single count of environmental policy measures – such as the environmental policy stringency (EPS) index as in Albrizio et al. (2017) and Botta and Kozluk (2014). Other studies have looked at how environmental policy affects markets by conducting event studies around the implementation dates of specific policies (Sato et al., 2020; Kruse et al., 2020; Mukanjari and Sterner, 2018; Barnett, 2020).

Second, our work builds upon the rapidly growing field in economics linking textual analysis of news to economic outcomes. An overview of the field and of the literature can be found in Gentzkow and Shapiro (2010). Closely related to our work, Baker et al. (2016) introduced the methodology to build indicators of economic policy uncertainty using news articles. Also related, Bybee et al. (2020) conduct a topic model analysis of business news published in the *Wall Street Journal* over 1984-2017 and find that specific news topics – for instance on recessions – have a significant predictive power on future output and employment. In contrast with developments in macroeconomics, applications of textual analysis in environmental economics remain limited (Dugoua et al., 2021; Baylis, 2020). A few papers link pollution news (Hamilton, 1995; Dasgupta et al., 2001) or investors' sentiment on renewable energy (Reboredo and Ugolini, 2018; Song et al., 2019) to stock prices. More recently, several studies in the financial literature have used text-mining techniques to measure climate risks (Sautner et al., 2020; Kölbel et al., 2020; Engle et al., 2020). Engle et al. (2020) for instance collect climate change news from the *Wall Street Journal* to provide a measure of climate risks as perceived by investors. Outside economics, a few studies have similarly investigated climate change news applying automated textual analysis and topic modeling of newspapers (Bohr, 2020; Keller et al., 2020; Dahal et al., 2019). So far, most of this literature using text analytics focuses on the salience of environmental problems (e.g. global warming) in public opinion. By contrast, our focus is solely on the salience of environmental *policy*, as a solution to environmental problems. In addition, we provide a more sophisticated methodology to identify and classify relevant news using automated text-mining techniques, combining both supervised and unsupervised machine learning algorithms.

The paper is organized as follows. Section 3.2 describes the data and methodology used to construct our news-based index of US environmental policy, as well as various descriptives and validity checks. Section 3.3 presents additional measures, namely a general measure of sentiment and various topic-specific indices, which can be directly derived from our general index. Section 3.4 examines the relationship between our environmental policy index and proxies for investments in clean markets, namely cleantech venture capital deals and stock returns – both at the firm-level in panel regressions and the aggregate level in VAR settings. Section 3.5 concludes.

3.2 A News-based Index of Environmental Policy

3.2.1 Developing a news-based environmental policy index

Our analysis starts from a set of 15 million articles from the archives of ten US newspapers available via automated access through Dow Jones' Factiva platform over the 1981-2019 period. We obtain access to the following newspapers: *New York Times, Washington Post, Wall Street Journal, Houston Chronicle, Dallas Morning News, San Francisco Chronicle, Boston Herald, Tampa Bay Times, San Jose Mercury News and San Diego Union Tribune.* Table C.1 in Appendix C provides detailed starting dates of the archive and additional statistics on the distribution of articles across newspapers.

As news about environmental regulations are relatively rare, manual retrieval of environmental policy articles is highly challenging. Hence, we first reduce our initial sample by selecting articles that contain keywords related to both 'climate change and the environment' and 'policy and regulations' within a proximity of 40 characters. Our choice of keywords in the category 'climate change and the environment' includes terms related to clean technologies[8], which cover energy generation, energy storage, energy efficiency, lighting, pollution (air, water, land), transportation (batteries, electric vehicles, clean fuels, etc.), recycling and waste and have been cross-checked against a broad definition of climate and environmental keywords from Climate Tagger.[9] The set of keywords on policy and regulations is borrowed from Baker et al. (2016) to which we add specific terms related to environmental policy (e.g. 'feed-in tariff').[10]

At this stage, our search strategy is extremely broad, as we want to avoid missing out on

[8]We used the website www.cleantech.org for our main definition of clean technologies. This definition excludes topics of conservation, fisheries, forestry and biodiversity and natural resources. These issues are a priori less relevant for clean investments and markets.

[9]www.climatetagger.net

[10]The full set of keywords and Boolean operators are available upon request.

potentially relevant articles (type-II error) by imposing too much structure. We obtain a set of 459,000 articles resulting from our query. For now, a large set of these articles are irrelevant as many general terms such as 'environment' or 'climate' are used in a myriad of contexts not applicable to environmental and climate regulations (type-I error). These articles refer for instance to 'tax policies to improve the business climate', 'agreement for changing the political environment', 'sustainable plans', etc.

Training sets, text pre-processing and classification

Our objective is then to correctly identify articles on environmental policy within our set of 459,000 articles. To do so, we use supervised machine learning which has two attractive features. First, it circumvents the need for a manual labeling of the entire corpus. Second, machine learning imposes only a minimum level of structure on what constitutes a relevant article (i.e., an article about environmental policy), in contrast with the more restrictive use of a combination of keywords.

The first step when using a supervised learning approach is to build manually annotated training sets which inform the algorithm about the content of environmental policy articles. We start by reading a large number of newspaper articles in order to develop a codebook defining criteria to classify environmental policy articles.[11] We then randomly select sets of articles to build two training sets: 1) an initial set of 995 articles from the *New York Times*, due to its high editorial quality and because its archive could be crawled early on in the process, 2) a set of 1,469 articles from our whole sample of newspapers, which better reflects the diversity of environmental policy articles. Three annotators then separately review overlapping sets of articles and manually assess whether or not a given article discusses environmental policy, guided by our codebook classification. About 20% of articles in the training set are classified as relevant for environmental policy by the annotators.

As a second step, we apply a set of standard text pre-processing steps to our corpus of 459,000 articles (removing very short articles, removing html tags, numbers and punctuation, lower-casing all words, stop-words filtering and lemmatization). Following standard approaches in computational linguistics, we convert articles into numerical vectors of unigram and bigram frequencies using a 'bag-of-words' approach, i.e. disregarding grammar and word order. We then construct a standard term-frequency inverse-document frequency (tf-idf) matrix in which less weight is given to words that occur either very frequently in the corpus or are barely used in the articles where they appear, because these are less informative than other words.[12]

In a third step, we use our training sets and the tf-idf matrix as inputs for a support vector

[11] Our full codebook is available on the website dedicated to our index: https://www.financingcleantech.com/envp-index

[12] More specifically, given a term-frequency matrix $tf(n,m)$, such that n is the number of articles and m the number of words, each term-frequency count is multiplied by the inverse document frequency. The Inverse document frequency ($idf_{j,n}$) is given by $log\left(\frac{N}{n_j}\right)$, where N is the total number of documents and n_j is the total number of documents containing j.

machine (SVM). SVM is a predictive data-classification algorithm which learns from the training set how to assign labels (i.e., environmental policy or not) to articles based on their most distinctive text features. We provide further details on the SVM algorithm, its parametrization and cross-validation in Appendix C.
We find that when the SVM model classifies an article as pertaining to environmental policy, it is correct 78 percent of the time (i.e., a precision of 0.78). This is good considering that, because only 22 percent of the articles in the training sets were labelled as relevant, classifying the articles at random would yield a precision of only 22 percent. Moreover, even with a codebook, deciding whether an article is about environmental policy or not requires a subjective judgement on the part of the annotators. If humans cannot perfectly identify relevant articles, we cannot expect our algorithm to do so either. In addition, the algorithm successfully retrieves more than two thirds of the relevant articles (i.e., a recall of 0.67).[13] This is satisfactory given that we are mostly interested in identifying trends in the relative volume of environmental policy news over time, rather than the exact universe of all relevant articles.

Finally, we apply the prediction rule produced by the algorithm on the whole sample. For each of our 459,000 articles, we input its respective word content to the algorithm, which then predicts whether the article belongs to the 'environmental policy' category or not. Our classifier identifies 80,045 relevant articles. Hence, less than 20 percent of all articles from our broad query end up being relevant in our final corpus, which is in line with insights from our manual labeling exercise.

Descriptive statistics

Table 3.1 displays the features that have the highest weight in predicting whether an article talks about environmental policy or not, according to our classifier. All features are those that one would expect to find in an article about environmental policy. They are a mix of both environmental (i.e. 'energy', 'emissions', 'environmental' or 'climate change') and policy-related terms (i.e. 'obama', 'epa', 'standards', 'federal' or 'regulations'). The SVM algorithm assigns an SVM-score to each article, based on its probability of belonging to the 'environmental policy' category.

[13] Precision is the fraction of documents identified as relevant by the classifier that were indeed labelled as relevant by the annotators. Recall is the fraction of the relevant documents that are successfully retrieved by the classifier. The F1-score is 0.72 with a precision of 0.78 and a recall of 0.67.

Table 3.1: Top discriminating keywords for predicting our EnvP index according to the trained SVM classifier.

Word	Weight	Word	Weight	Word	Weight
energy	3.16	crist	1.34	volkswagen	1.09
emission	3.06	air	1.33	refrigerator	1.08
environmental	2.95	ethanol	1.32	utility	1.07
epa	2.24	global warming	1.32	cleanup	1.06
solar	2.17	coal	1.30	federal	1.05
obama	2.05	climate	1.26	car	1.00
clean	1.89	regulation	1.24	penalty	0.99
pollution	1.83	program	1.18	house	0.98
waste	1.67	renewable	1.17	bannon	0.98
warming	1.62	reef	1.15	bill	0.98
recycle	1.47	protection	1.14	mercury	0.97
power	1.45	climate change	1.12	electric	0.96
global	1.38	env. protection	1.10	gasoline	0.94
standard	1.36	clean air	1.10	environment	0.94

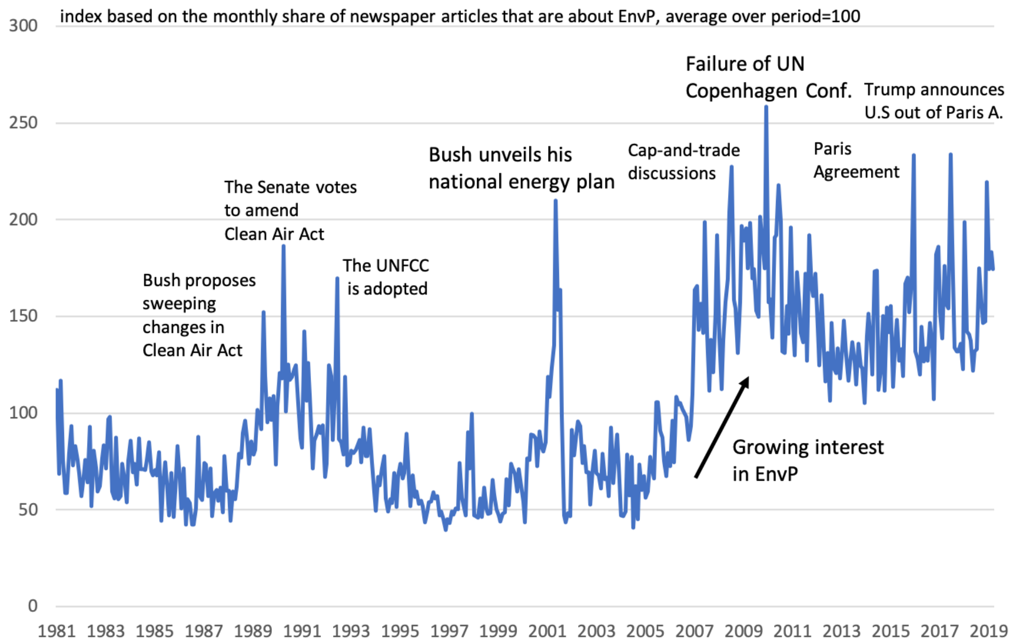
Table 3.2 reports excerpts of the five newspaper articles with the highest SVM score. All of these articles are extensively covering environmental policy issues, giving us confidence in our classifier. The first article titled *Environment — Handicapping the Environmental Gold Rush* is a special edition about the green transition and the crucial role that government policies will play in shaping the future of both dirty and clean energy. The second article *In Texas, clean energy set to boom*, describes the ongoing changes in the electricity sector in Texas and the impact of future air pollution regulations.

Table 3.2: Newspapers articles with the highest SVM-score

Title	Date	Score	Newspaper	Excerpt
Environment — Handicapping the Environmental Gold Rush	Oct 29, 2007	3.55	Wall Street Journal	"The green stampede is on. As a global economy powered by cheap fossil fuel comes under intense pressure to change, corporate executives are racing to stay ahead of the tectonic shift in their world. From Capitol Hill to California and Brussels to Beijing, multinational companies are stepping up their lobbying [...]"
In Texas, clean energy set to boom	Jan 10, 2016	3.54	Dallas Morning News	"While Texas has long been the top state for oil and gas, much more is going on here. In electricity, cleaner-burning natural gas plants are pushing out coal faster than in the rest of the nation, and that's before the next air pollution regulations kick in."
Obama Flies to the Nevada Desert to Promote Solar Energy	Aug 25, 2015	3.53	New York Times	"While promoting the benefits of all renewable energy, including wind power, the president focused largely on solar energy, part of an increasingly intense effort to counter global warming by instituting policies to reshape the nation's energy industry."
New rule targets pollution from coal	Aug 2, 2015	3.49	Washington Post	"The Obama administration will formally adopt an ambitious regulation for cutting greenhouse-gas pollution on Monday, requiring every state to reduce emissions from coal-burning power plants and putting the country on a course that could change the way millions of Americans get their electricity."
Environmentalists, Industry Air Differences on Pollution	Oct 17, 1999	3.48	Washington Post	"As a result, environmental groups are pressing states and Congress for specific environmental protections against increased pollution, financial incentives for energy efficiency and renewable energy, and federal pollution guidelines to be part of the overall deregulatory effort."

To construct our index, we count the monthly number of articles classified as *environmental policy* by our SVM algorithm. Since the total amount of news published in newspapers varies over time, we scale the monthly counts of environmental articles predicted by the total monthly volume of news articles in our ten newspapers. Figure 3.1 plots our environmental policy index (EnvP index). The index is normalized such that its average value over the 1981-2019 period is equal to 100. As shown in Figure 3.1, our index is able to capture both trends and spikes in recent media coverage of US environmental policy history. We observe more than a three-fold increase in media attention on environmental policy issues between 2006 and 2009. At the tail-end of this trend, our index identifies two important events which precipitated a fall in both media interest and political will to address climate change, namely 1) the parliamentary debates over the Waxman-Markey bill in April 2009 which failed to introduce a cap-and-trade system and 2) the UN Copenhagen Conference in December 2009 which ended on an unbridgeable North-South divide. Other salient events, such as the signature of the Paris Climate Agreement in December 2015 and President Trump's announcement of withdrawal from the agreement in June 2017, are labelled in Figure 3.1.

Figure 3.1: EnvP - An index of environmental policy, monthly share



3.2.2 Evaluating our news-based environmental policy index

In this section we provide further descriptives and validity checks of our news-based index of US environmental policy.

Comparison with other indices

We first compare our index to currently available newspaper-based environmental indices. In Figure 3.2 we compare our EnvP index with the Climate Change News index created by Engle et al. (2020) on a quarterly frequency.[14] The latter index is based on the occurrence of keywords related to climate change in articles published by the *Wall Street Journal* (WSJ). In contrast, our EnvP index aims to capture environmental *policy*, i.e. the policy solution to the climate change problem. In addition, we consider a much broader set of newspapers across several regions in the US, rather than the more international audience of the WSJ alone. As a result, policy terms (e.g., 'obama', 'epa', 'standard', 'regulation' or 'program') are well represented among the top features identified by our trained classifier in Table 3.1. By contrast, the most important features in the Climate Change News index are exclusively about climate change, the related science and international discussions.[15]

We find that the two series are positively correlated to one another with a correlation coefficient of 0.67 for the monthly series and 0.75 for the quarterly series.[16] Yet, there are notable differences between the two indices as shown in Figure 3.2. First, the Climate Change News index has less pronounced trends and reacts strongly to the UN Climate Talks (e.g. adoption of the UNFCCC in 1992, adoption Kyoto protocol in 1997, etc.). Second, our index appears to be better able to capture the large number of discussions on environmental policy on the US political agenda between 2008 and 2014, a period where climate change was not necessarily at the forefront (with one notable exception being the failure of the 2009 UN Climate Change Conference in Copenhagen). By contrast, many issues dealing with US domestic policies received much media attention in this period, such as the failed adoption by the U.S. Congress of a cap-and-trade bill in 2009 or the federal policy support for clean energy companies like Solyndra.[17]

An important question in our analysis is how our index relates to environmental policy stringency. We argue that the importance given to environmental policies in the news reflects how impactful these policies are in the economy. We therefore expect our EnvP index to be positively correlated with environmental policy stringency. Therefore, as an additional reality check, we compare our EnvP index (12-month moving average) to the OECD's Environmental Policy Stringency Index (EPS) for the United States in Figure 3.3. The EPS measures the extent to which a country puts an explicit or implicit price on polluting or environmentally harmful behaviour. We see that the indices co-move, with a correlation coefficient of 0.79 between 1990 and 2015. The EnvP index seems nonetheless more sensitive to one-off increases in

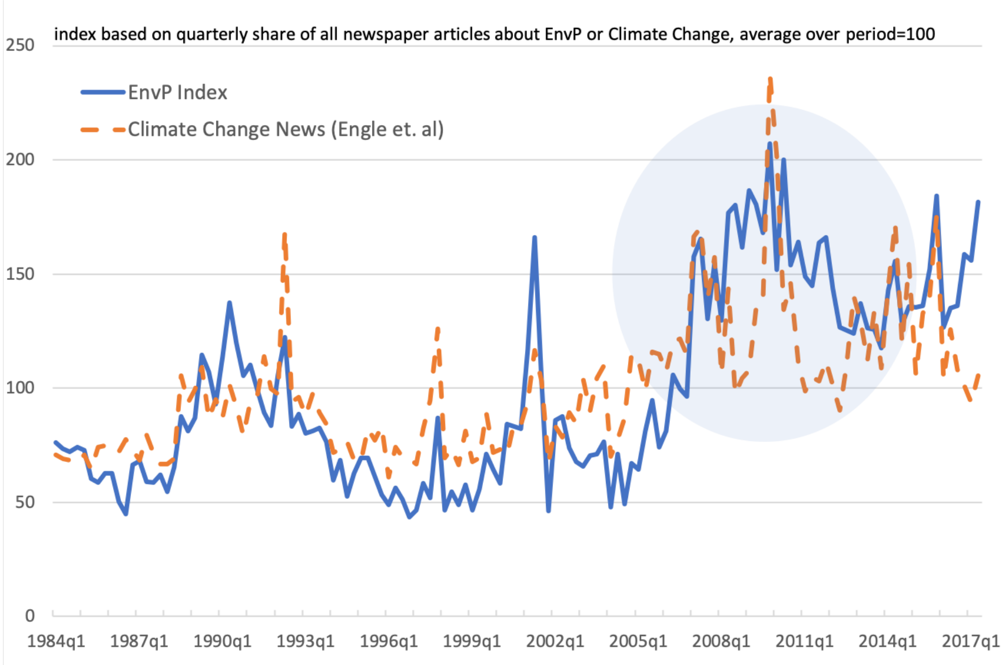
[14] We show them on a quarterly frequency rather than monthly to facilitate the comparison by smoothing out some of the variability.

[15] One policy term— 'protocol'—is present in their top features because of its use in international agreements.

[16] If we compare their index to our EnvP index based solely on the WSJ, a fairer comparison, the monthly correlation climbs up to 0.71. See Figure C.1 in Appendix C.

[17] Solyndra received a $535 million U.S. Department of Energy loan guarantee, and was the first recipient of a loan guarantee under Obama's economic stimulus program, the American Recovery and Reinvestment Act of 2009. The bankruptcy of Solyndra in 2011 has received a lot of attention in the media and was used by Obama's political opponents as an example of wasteful spending under the stimulus program.

Figure 3.2: Comparison with Engle et. al, quarterly share



media attention to environmental policy, such as the energy crisis of 2001.

Is environmental policy coverage partisan?

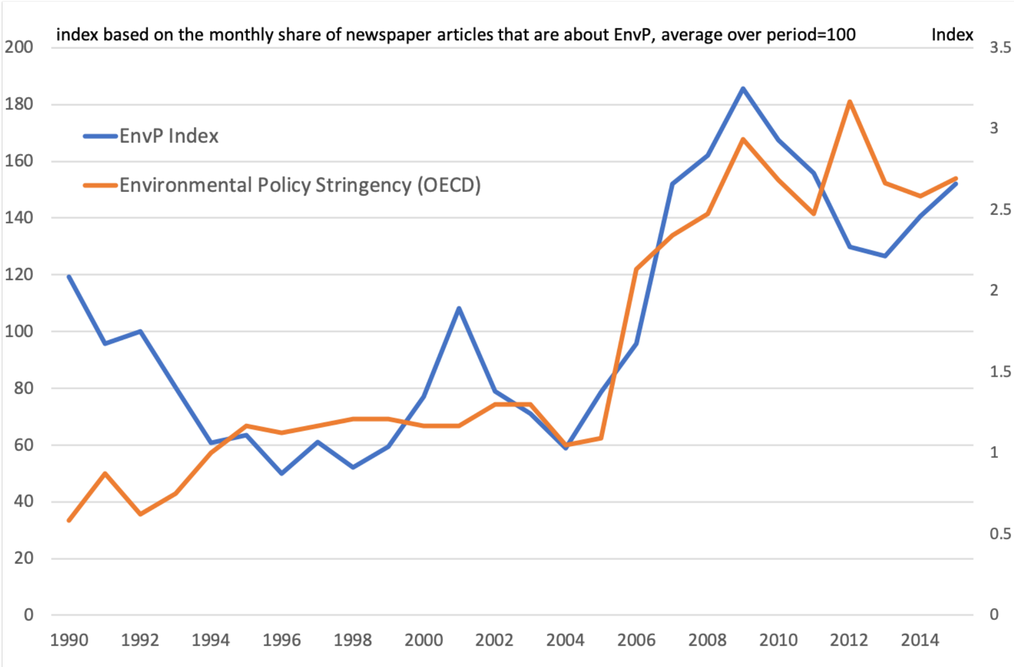
A newspaper-based measure of environmental policy should not be overly influenced by the political slant of the newspapers in our sample. To investigate this issue, we divide the newspapers in our sample into two groups based on whether they are more conservative or liberal-leaning.[18]

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

First, we find that 0.55% of articles in liberal-leaning newspapers are about environmental policy and 0.48% in the more conservative-leaning ones. We plot the EnvP indices produced by the liberal-leaning and conservative-leaning newspapers in Figure 3.4. The figure shows

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

Figure 3.3: Comparison with EPS, yearly share



that the coverage of environmental policy has followed the same trends in these two groups over the past four decades. There are only a few minor exceptions. Notably, liberal-leaning newspapers dedicate more space to environmental policy in the early-months of Trump's presidency than conservative-leaning newspapers. However, as our sample of newspapers is well balanced between liberal and conservative outlets, our general EnvP index averages out the differences. Overall, we observe that political slant does not seem to skew the coverage of environmental policy and is thus not a serious concern for our analysis.

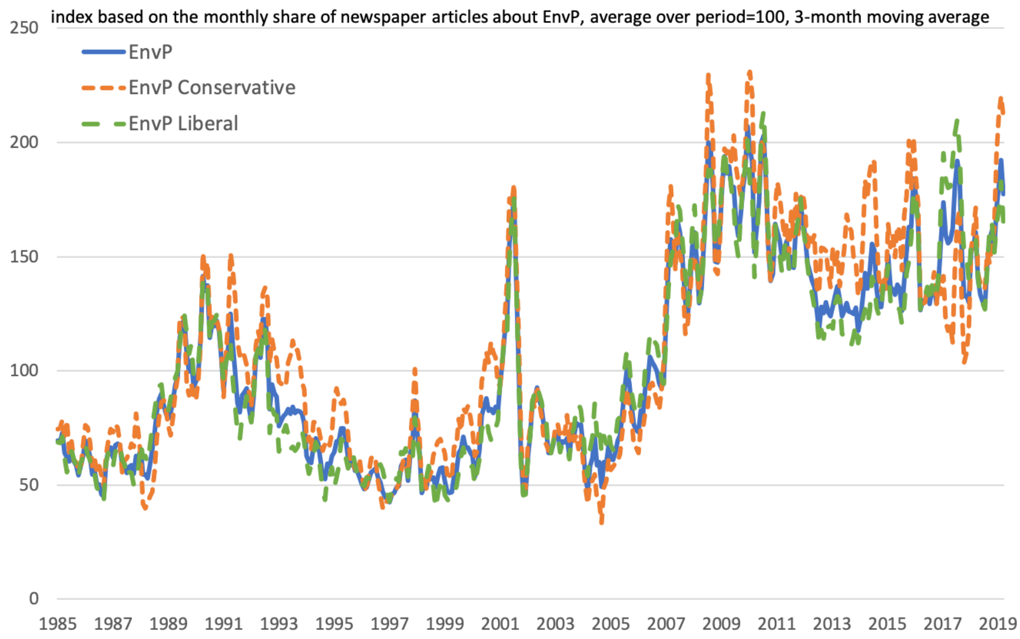
3.3 Additional Measures of Sentiment and Topic-specific Indexes

In this section, we introduce two additional types of measures that can be extracted from our main index of environmental policy, namely: 1) a sentiment index and 2) topic-specific indices. These measures help illustrate the vast amount of information that can be extracted from newspapers and help us refine some of our empirical analysis in Section 3.4.

3.3.1 Sentiment analysis

News about environmental policy may be either positive, negative or neutral. To assess the polarity of our index, we conduct a sentiment analysis following Consoli et al. (2021). The authors develop a novel fine-grained aspect-based methodology that allows for identifying topic specific sentiment around certain keywords, as opposed to merely identifying the general

Figure 3.4: EnvP according to liberal and conservative media

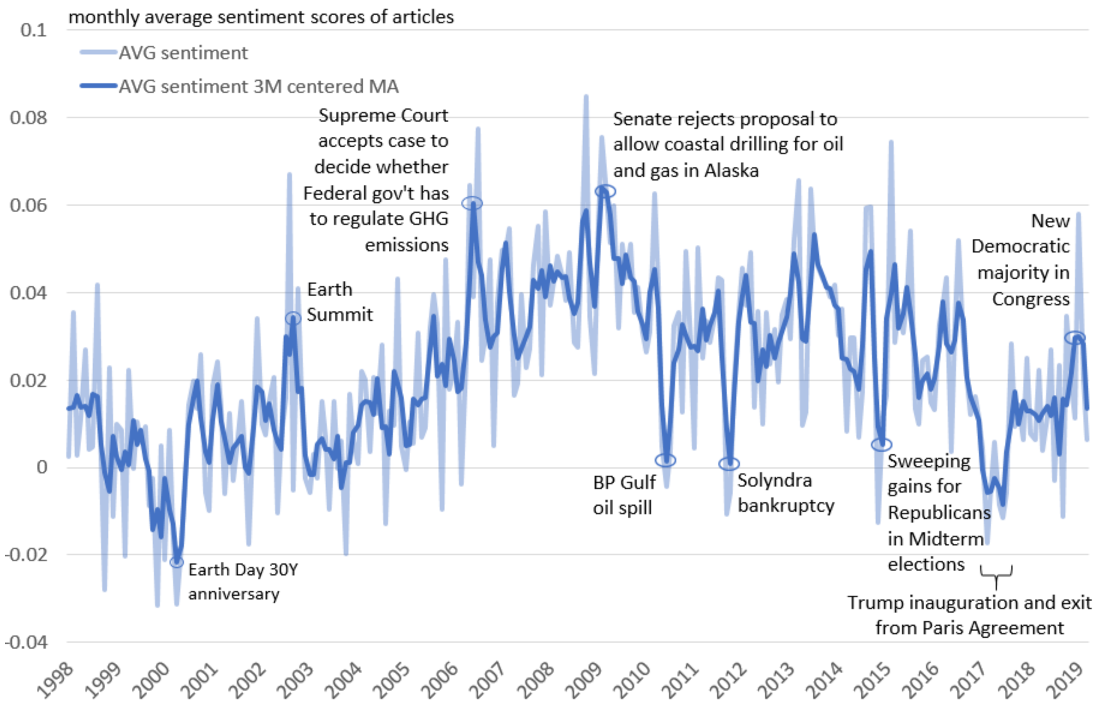


sentiment of a given sentence. The advantage is that, by selecting suitable policy keywords, we make sure to pick up the sentiment actually pertaining to policy, not to any confounding sources of sentiment such as financial markets. A central element in sentiment analysis is the dictionary which assigns each word in the lexicon a sentiment score. In Consoli et al. (2021), the lexicon base is optimized for economic and financial texts.[19]

As it is our aim to measure sentiment pertaining to environmental policy, the algorithm, searches around the following terms: ‘federal’, ‘state’, ‘court’, ‘treaty’, ‘summit’, ‘political’, ‘administration’, ‘talks’, ‘policy’, ‘congress’, ‘epa’, ‘senate’, ‘regulation’, ‘rule’, ‘penalty’, ‘program’, ‘house’, ‘bill’, ‘protection’, ‘legislation’, ‘standard’. The term set was selected from the list of keywords most important for our EnvP classifier as we can be sure that these appear in most of our EnvP articles. For each article, the algorithm identifies sentences where one of the words specified above appears and searches around it to generate a sentiment score per relevant instance. This yields multiple scores per article which all vary between -1 and +1, with zero

[19] We also considered the dictionary-based approach by Loughran and Mcdonald (2011). The downside of this approach for this application is that, as mentioned in Consoli et al. (2021), their dictionary is based on 10-K filings and not on newspaper articles. Moreover, their list of terms for negative sentiment is much longer than that of positive sentiment which inevitably affects the level of sentiment as picked up by the index. However, despite the difference in average sentiment, the trends in the sentiment indices based on Consoli et al. (2021) and Loughran and Mcdonald (2011) are roughly similar with a correlation of 0.46 for the raw index and 0.71 for the six-month moving average. This gives us confidence that the general trends in sentiment are not driven by any methodological particularity. Furthermore, Consoli et al. (2021) show that their methodology outperforms Loughran and Mcdonald (2011) and other common methods in the literature when comparing the predicted labels with a human-annotated sample of texts.

Figure 3.5: EnvP sentiment index



representing neutral sentiment. Next, we compute the average sentiment score for each article which leaves us with one sentiment score per article.

Figure 3.5 displays our EnvP average sentiment index. Sentiment has been fluctuating around the neutral cutoff of zero until the end of 2002 after which monthly sentiment scores tend to remain positive. The uptick in sentiment in the 2000s is likely driven by a general increase in public discussions on climate change (e.g. through Al Gore's movie 'An Inconvenient Truth', Jacobsen (2011)) and coincided with a strengthening of environmental policy in the US with favourable reporting in the news. Other salient positive and negative events are labelled in Figure 3.5. Later dips in sentiment include the 2014 US elections which led to sweeping gains for the Republicans threatening to thwart Obama's climate policy agenda, corresponding to the largest two-month dip in sentiment across the whole sample. Finally, as expected, we also find a major dip in sentiment in June 2017 when Trump announced the exit from the Paris agreement.

3.3.2 Topic-specific indices

Finally, we show that our index captures a lot of fine-grained information on specific topics. We apply topic modeling, an unsupervised machine learning approach, to illustrate how our index can be decomposed to identify specific environmental policy topics.[20] As an example, Green New Deal policies may include provisions specific to sub-topics on ’automobile emissions’ or ’renewable energy’. Unsupervised learning approaches can help discover implicit patterns in the data without researchers imposing any specific structure (such as keywords or a training set). This technique identifies re-occurring word patterns to infer a given number of topics within our corpus of articles.

As a first step, and to limit the number of unique terms included in our analysis, we build a tf-idf matrix of the whole sample of unigrams, bigrams and trigrams included in our 80,045 environmental policy articles and select the 20,000 with the highest tf-idf score.

We then apply Latent Dirichlet Allocation (LDA) topic modeling developed by Blei et al. (2003) and already successfully applied in the economics literature (Hansen and McMahon, 2016; Hansen et al., 2018; Bybee et al., 2020). LDA is a statistical model that views each document as a collection of topics and each topic as a collection of keywords. A given keyword can be attributed to different topics with varying proportions and, likewise, an article can be 80 percent about ‘automobile emissions’ and 20 percent about ‘renewable energy’. We provide detailed information on the LDA algorithm and our methodological choices regarding the number of topics in Appendix C. We identify a topic model with 26 topics to be optimal.[21] In the remainder of the analysis, we focus on the two topics of ’renewable energy policy’ and ’international climate negotiations’ for illustrative purposes. An in-depth discussion and analysis of the various topics is provided in a companion paper (Noailly et al., 2022).

In order to interpret the topics uncovered by the LDA, we look at the most prevalent words per topic. Figure 3.6 displays the keywords for two chosen topics using word clouds. The size of a word within a cloud corresponds to its probability of occurring within the topic. The word cloud on the left of Figure 3.6 is composed of terms such as *energy, solar, wind, power, renewable, electricity, credit, etc.* We label this topic as ‘renewable energy policy’. Similarly, the word cloud on the right of Figure 3.6 of the combination of words: *united, country, agreement, united_state, world, international, environment, etc.* We label this topic as ‘international climate negotiations’.

To construct topic-specific indices, we count the number of articles attributed to a given topic and scale it by the total volume of newspaper articles in our sample. As before, we scale the indices so that their average corresponds to 100 over the 1981-2019 period. Figures 3.7 and 3.8 plot the evolution of the index for the sub-topic of renewable energy policy and international climate negotiation, respectively.

[20] Because our index captures a lot of fine-grained information on US domestic policies, our index likely captures a much richer set of topics than the index developed by Engle et al. (2020) constructed using a fixed set of keywords on climate change vocabulary.

[21] The list of topics is provided in Table C.2 of Appendix C. None of the topics captured by our model are completely unrelated to environmental policy, which is further evidence of the relevant content of our 80,045 EnvP articles. Using our topic model, we can see how the environmental policy discourse has evolved over time. Figure C.5 shows how some of the most important topics have evolved, with ‘renewable energy policy’ and ‘climate change science’ in particular gaining increasing media attention since the 1990s.

Figure 3.6: Renewable energy policy & International climate negotiations Word Clouds

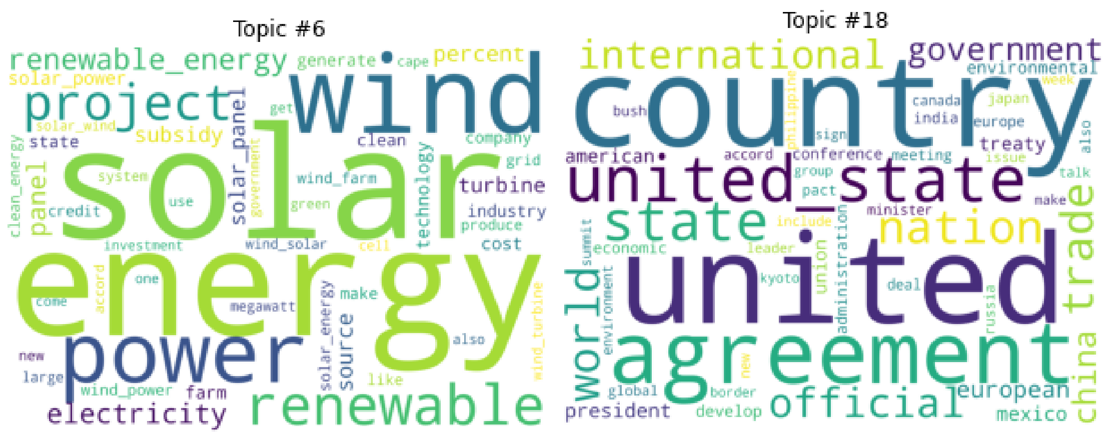


Figure 3.7: Index - Renewable energy policy

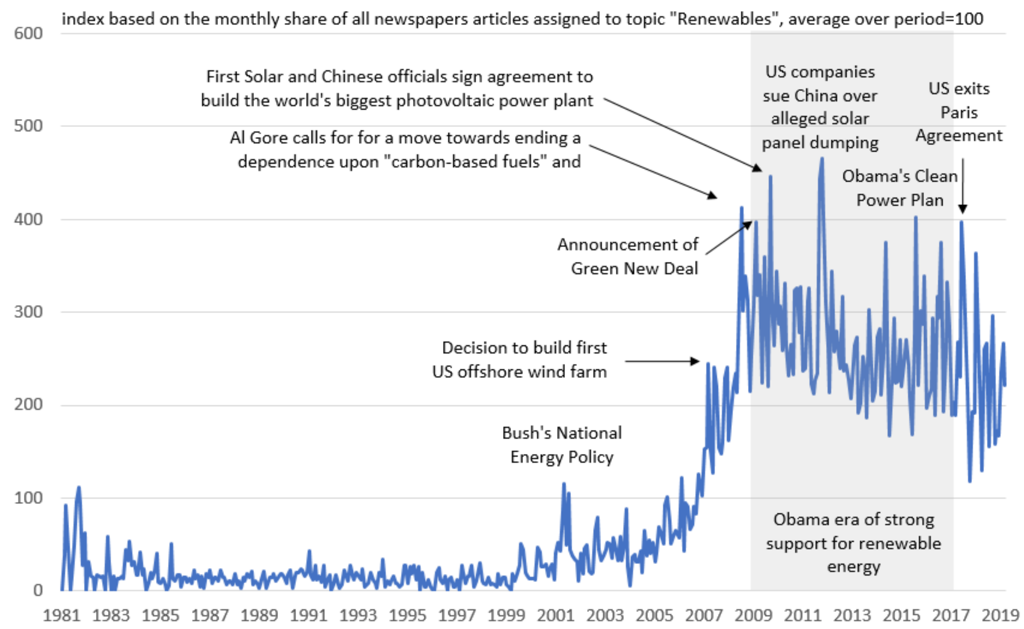
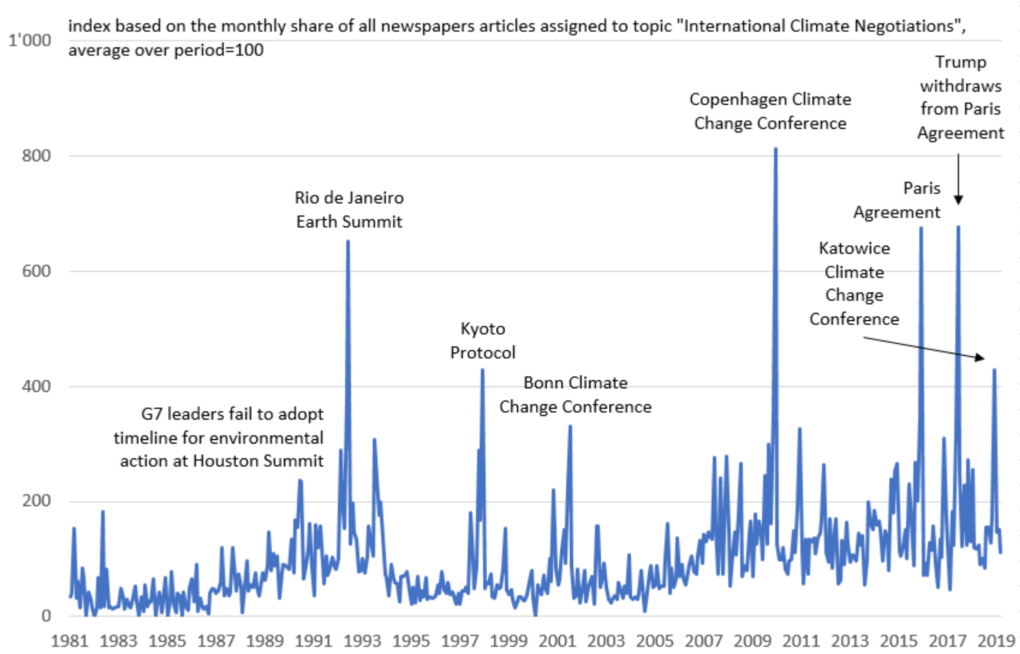


Figure 3.7 shows that policy discussions on renewable energy rise significantly after 2005, which corresponds to the period of implementation and strengthening of renewable portfolio standards in many US states, particularly during the Obama presidency. The announcements of the Green New Deal in February 2009 and the Clean Power Plan in August 2015 aiming to reduce carbon dioxide emissions from electrical power generation by 32 percent by 2030 relative to 2005 levels marked key milestones toward this agenda. Other specific issues reported in news articles are indicated directly on Figure 3.7.

Figure 3.8 shows that our international climate negotiations index correctly captures discus-

Figure 3.8: Index - International climate negotiations



sions around major international climate summits. It identifies events such as the Rio Earth Summit in June 1992, the Kyoto Protocol in December 1997 and the Bonn Climate Change Conference in July 2001 which was the first international meeting after President Bush had exited the Kyoto Protocol in March 2001. In more recent years, the Copenhagen Climate Change Conference in December 2009, the Paris Agreement in December 2015 and the COP24 in Katowice are all picked up as salient events by our index.

Finally, beyond this selective set of illustrations, we believe that there are many additional powerful applications of our EnvP index. Our index and topic model can for instance easily be combined with keywords to identify a more fine-grained set of policies. We provide in Figures C.2 and C.3 in Appendix C further illustrations of a news-based index for the state of California and of an index depicting the environmental policy coverage around ExxonMobil.

3.4 Environmental Policy and Investments in Clean Technologies

We now turn to the central part of our analysis, documenting the meaningful association between our EnvP index and clean markets. Conceptually, we expect that a rise in the volume of news on environmental policy signals growing policy support and thus increasing opportunities for clean products, technologies and firms. In other words, we expect a rise in our EnvP index to be associated with an increase in investments in clean markets. Conversely, an increase in our EnvP index could signal vanishing opportunities for dirty markets, so we may also expect our EnvP index to be negatively associated with investments in dirty and polluting products and firms. In our empirical analysis, we consider two proxies for investments, namely venture capital finance and stock markets valuation, and conduct our analysis both at the micro level using firm-level estimations and at the aggregate level using VAR models. While the first approach provides better suggestive evidence towards causality, the second approach provides insights on the dynamic relationship between the salience of environmental policy and investments in clean energy markets at the sectoral level and might potentially capture additional channels (e.g. entry and exit).

3.4.1 Environmental Policy and Firm-Level Clean Investment Decisions

Establishing causality on the impact of environmental policy news on economic activity is challenging. First, media reporting on environmental regulations may respond to expectations of future clean market growth and technological advancements. Investors might also anticipate policy news which could lead to a downward bias in our estimate. Second, both media attention and clean investments may be affected by additional omitted variables, such as growing environmental awareness, which is likely to raise both environmental policy news and investments in clean technologies, leading to an upward bias in our estimate. Past studies have evaluated the impact of policy by using event study methodologies around a discrete unanticipated policy change announced in the media. Event studies give more assurance of causal effects but are limited to a specific timeline of events, while the advantage of our index is to measure a continuous tracking of policies. Our identification strategy relates to a difference-in-difference approach, where we consider various firms that inherently differ in their exposure to environmental regulations. We show in firm-level estimations how financial investments in startups and firms most exposed to environmental policy, as defined by their sector of activity or emission levels, are associated to a higher salience of environmental policy, as measured by our EnvP index.

VC investments across industries

We first examine how our EnvP index is associated with the probability that a startup will receive VC funding. So far, the literature on the determinants of VC funding for cleantech startups is quite limited (Nanda et al., 2015; Popp et al., 2020). Yet, one of the main advantage of VC data for our purpose is that information on deals are available at a high frequency (i.e., daily).

We obtain data on VC funding rounds between January 1998 and March 2019 for US startups from the Crunchbase database and aggregate these funding rounds into a firm-quarter panel dataset.[22] We also extract firms' industry and founding date as well as all the information

[22] We focus on series A to J financing, involving firms founded after 1985. This represents around 75,000 different funding rounds. Due to the panel nature of our dataset, we observe startups over long period of times and therefore should avoid including inactive startups. These inactive startups can either have gone bankrupt or not be in need of early-stage financing anymore and as such have a probability equal to zero to receive VC funding. We therefore classify any firm that fails to secure new financing within the three-year time span after its last round of financing

related to the funding rounds (i.e., date, amount, series) from Crunchbase. Finally, we include GDP data from the U.S. Bureau of Economic Analysis, the Federal Reserve effective funds rate and the West Texas Intermediate crude oil spot prices. Excluding firm-quarter observations containing any missing information–including those where the firm is classified as inactive–we obtain 1,056,221 firm-quarter observations on 35,637 unique startup firms.

Our identification strategy differentiates startups by their exposure to environmental policy: more precisely, we expect to find a positive impact on the probability of VC funding for startups classified in cleantech industries, while we expect no significant impact for other startups. Cleantech startups belong to Crunchbase's 'Sustainability' industry group and represent 4% of overall VC deals, while clean energy startups in clean energy, battery, renewable energy, wind energy, energy storage and solar industries represent only 2.4% of all VC deals.We estimate whether startups that are classified as cleantech or clean energy are significantly more responsive to our EnvP index than startups in other sectors using ordinary least squares (OLS) as follows:[23]

$$\begin{aligned} VC_{i,t+1} =& \alpha + \beta_1 EnvP_t + \beta_2 EnvP_t \cdot Cleantech_i + \beta_3 Controls_{i,t} + \\ & \beta_4 TimeTrend_t + \gamma_{quarter/year/industry/state/series} + \epsilon_{i,t} \end{aligned} \tag{3.1}$$

where i indexes the firm, t the quarter. We use two different measures of VC investments as our dependent variable: a funding dummy, $Funded$, and the logarithm of the total amount of funding a startup receives during a quarter, $Amount$, conditional on $Funded = 1$. β_1 and β_2 are the coefficients of our two main variables of interest. β_1 captures the impact of $EnvP_t$ –which takes two forms, either our EnvP Index or our EnvP-RE sub-index– on non-cleantech (clean energy) startups. β_2 on the other hand captures the particular impact of our EnvP variable on cleantech (clean energy) startups.

We control for the following variables that could be confounding our results. First, our EnvP index is likely affected by business cycles effects, as environmental concerns might take a backseat role during a crisis. We therefore control for economic activity and capital availability by including the annual growth of U.S. GDP and the Federal Reserve effective funds rate. Second, we include the log of the oil price as it is both an important actor in the environmental policy debate and actual investment decisions. In some specifications we include the output of our sentiment analysis on our EnvP index. This allows us to control for the positive (or negative) content of the news. We also include a set of variables and fixed effects to absorb variation that is unrelated to environmental policy but may nonetheless affect our results, including, firm i's age in quarter t –set as missing before founding date and if it is inactive– as well as a time trend, and in some specifications an industry time trend. We also use firm, quarter, year and series funding round fixed effects.[24] The firm fixed effect control

as inactive after this three-year mark.

[23] Our results are robust to using Probit, see Table C.4 in Appendix C

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

for firm-level unobservables such as quality. The quarter fixed effects are used to account for seasonality in the data.[25] The other fixed effects also allow us to control for unobserved variables common to all startups in a given year or funding round. Finally, we cluster standard errors at the startup firm level to correct for potential serial correlations in the error term.

We first focus on the relationship between our EnvP index and VC investment in cleantech using Equation (3.1). We present the regression results in Table 3.3, first using the probability of getting funded in the next quarter (Q + 1) as the dependent variable. Using Column (1) we can see that while a rise in our EnvP index is associated with a marginally lower chance of receiving funding for the average startup, it has a strong positive relationship with cleantech startups' probability of receiving funding in the next quarter. To illustrate the size of the effect, a doubling of environmental policy media coverage from one quarter to the next is associated with an increase in the probability of receiving funding of 1.4 percentage points.[26] While this might seem like a small increase, the average probability that a cleantech startup will be funded next quarter in our sample is only 6.2 percent. Therefore a doubling of environmental policy media coverage is actually associated with a 23 percent increase in a cleantech startup's probability of receiving funding next quarter. Column (2) shows that news with a positively-toned sentiment are associated with more VC deals in cleantech. After correcting for sentiment, the EnvP index continues to be significant, indicating that both the volume and sentiment of articles matter to investors. In Column (3), we see that our result holds when our EnvP index is orthogonalized to the Climate Change News index by Engle et al. (2020). Moreover, it shows that, keeping the EnvP index constant, the Climate Change News index has no particular relationship with cleantech VC investments. This could be evidence that what matters for VC investments in cleantech is media attention to climate change policies and not necessarily to climate change itself. Columns (4) and (5) use the natural logarithm of the amount received in dollars, given that they received funding next quarter, as the dependent variable. We find that a one percent increase in our EnvP index is associated with a 0.6 percent increase in the amount received by cleantech startups. This remains robust when controlling for the sentiment index.

Overall, the relationship between the EnvP index and VC investments in non-cleantech startups is in line with our expectations. It is either insignificant (Columns (4) and (5)) or negative (Columns (1) through (3)). The negative coefficients in the first three columns could indicate that rises in EnvP cause VCs to allocate a higher share of their funding to cleantech startups, at the expense of non-cleantech startups. In any case, all the coefficients highlight the fact that the EnvP index is disproportionately associated with cleantech firms.

As an additional robustness check, we consider how our sub-topic index of media attention on renewable energy policies (EnvP-RE) relates to the financing of startups active in renewable energy and fossil fuels industries. We expect renewable energy startups to be more affected by EnvP-RE news than other startups. However not all energy startups should become more

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

[26]We obtain this number by doing the following calculation $(-0.00538 + 0.0253) \cdot 0.693$, given that a doubling in a logged variable implies an increase by 0.693.

Table 3.3: Relationship between EnvP media coverage and VC investments in cleantech

	(1) Funded (Q+1)	(2) Funded (Q+1)	(3) Funded (Q+1)	(4) Amount (Q+1)	(5) Amount (Q+1)
Log EnvP Index	-0.00538***	-0.0110***	-0.00652***	-0.0168	-0.0188
	(0.00194)	(0.00194)	(0.00217)	(0.0484)	(0.0491)
Log EnvP x Cleantech	0.0253***	0.0212***	0.0228***	0.467***	0.379***
	(0.00479)	(0.00485)	(0.00547)	(0.147)	(0.142)
Log Sentiment Index		-0.00901***			-0.0220
		(0.000659)			(0.0136)
Log Sentiment x Cleantech		0.00936***			0.278***
		(0.00216)			(0.0628)
Log Climate Change News Index			-0.00759***		
			(0.00204)		
Log Climate Change News x Cleantech			0.00514		
			(0.00645)		
Firm FE	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Industry-time trend	Yes	Yes	Yes	Yes	Yes
Series FE	Yes	Yes	Yes	Yes	Yes
Observations	1056221	1056221	935517	57319	57319
Firms	35637	35637	34218	28297	28297
R^2	0.006	0.006	0.007	0.118	0.119

The table presents results of an OLS regression. The sample period is January 1998 and March 2019. The dependent variable in Columns (1) to (3) is a dummy variable that indicates whether firm *i* received VC funding next quarter. In Columns (4) and (5), the dependent variable is the logarithm of the amount received, conditional on having received funding. Controls include age, oil price, a time trend, GDP and the fed fund rate. Standard errors are clustered at the company level. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

attractive to VC investors when media attention on renewable energy policy is high. VC funding of fossil fuel startups should therefore have no, or a negative, relationship with the EnvP-RE. Table 3.4 displays our results. Results on the interaction terms show that a rise in EnvP-RE media coverage is associated with both a higher probability for clean-energy startups to secure funding and a higher amount per funding. At the same time, the EnvP-RE index has no significant relationship with VC investments in fossil fuels startups.

Firm-level stock returns

Next, we examine how our EnvP index relates to firm-level stock returns in panel estimations, drawing on the emerging literature looking at how environmental policy signals are reflected in firms stock valuations (Kruse et al., 2020; Mukanjari and Sterner, 2018; Barnett, 2020).

We start our analysis by collecting monthly total return indices for a sample of around 1400

Table 3.4: Relationship between the renewable policy index and VC investments in clean and dirty energy

	(1) Funded (Q+1)	(2) Amount (Q+1)
Log EnvP-RE index	0.00657***	-0.00378
	(0.00119)	(0.0293)
Log RE x Renewables	0.0134***	0.616***
	(0.00312)	(0.106)
Log RE x Fossil Fuels	-0.00497	0.00434
	(0.00488)	(0.121)
Firm FE	Yes	Yes
Quarter FE	Yes	Yes
Year FE	Yes	Yes
Controls	Yes	Yes
Industry-time trend	Yes	Yes
Series FE	Yes	Yes
Observations	1056221	57319
Firms	35637	28297
R^2	0.006	0.119

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

firms across various industries listed on the US stock exchange from January 2004 to March 2019 from Datastream. We also extract the monthly safe interest rate from the website of Kenneth French[27] and compute monthly continuously compounded log returns at the firm level as $r_{i,t} = \ln\left(\frac{p_{i,t}}{p_{i,t-1}}\right)$. We use the excess returns above the safe rate r_f, i.e. $r^e_{i,t} = r_{i,t} - r_f$, as our dependent variable. In the estimation, market returns are controlled for on the right-hand side by five market risk factors (Fama and French, 2015).

When working with stock price data, we may be particularly worried that investors may anticipate and react very quickly to movements in our EnvP index. To mitigate this, we follow Brogaard and Detzel (2015) in considering 'innovations' in our EnvP index (i.e. its unanticipated component) by extracting the residuals from an AR(7) model of our monthly series of EnvP as follows:

[27] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

$$\epsilon_t^{EnvP} = EnvP_t - \left(\hat{\phi}_0 + \sum_{k=1}^{7} \hat{\phi}_t * EnvP_{t-k}\right) \tag{3.2}$$

Standard tests confirm that this series is white noise and has no autocorrelation.[28] We standardize this measure to have a mean of zero and a unit standard deviation.[29]

Our identification strategy differentiates firms according to their exposure to environmental policy. To proxy for such exposure, we use firm-level scope 1 CO_2 emissions[30] collected from publically disclosed data sources at annual frequency from S&P Trucost Limited. As environmental regulations gain prominence in the news, we expect investors to divest from high-emission firms, leading to lower stock returns. Since CO_2 emissions and thus firm exposure to environmental policy may endogenously respond to current or anticipated environmental policy, we use the (standardized) fixed mean of CO_2 emissions over the sample period in our baseline estimation. We consider the following panel estimation:

$$r^e_{i,j,t=m} = \alpha + \beta_1 \epsilon^{EnvP}_{t=m} + \beta_2 \text{CO2 Emissions}_{i,t=y} + \beta_3 \text{CO2 Emissions}_{i,t=y} * \epsilon^{EnvP}_{t=m} + \tag{3.3}$$

$$\beta_4 \text{Risk Factors}_{t=m} + \beta_5 \text{Firm controls}_{i,t=y} + \beta_6 \text{Time Trend}_{j,t=y} + \gamma_{\text{i}} + \tag{3.4}$$

$$\varepsilon_{i,t=m} \tag{3.5}$$

where i, j, t indicate firm, industry and time (with m denoting month and y denoting year), respectively. CO2 Emissions$_{i,t}$ proxies firm-level environmental policy exposure by the fixed mean of CO_2 emissions over the period. $\epsilon^{EnvP}_{t=m}$ represents the monthly EnvP innovations. Risk Factors$_t$ is a vector containing the monthly market risk factors MKTRF, SMB, HML, RMW and CMA from the 5-factor Fama-French asset pricing model (Fama and French, 2015) obtained from the website of Kenneth French.[31] In addition, $X_{i,t}$ is a vector of firm-specific characteristics, namely (i) firm size as log(market cap), (ii) a measure of firm profitability as log(return on assets), (iii) a measure of firm leverage as log(total debt/total equity) as well as (iv) log(dividends per share). Table C.5 in Appendix C provides summary statistics of all variables used in the analysis. Finally, we include an industry-year time trend in all our specification to control for time-varying factors specific to industries, such as technological progress, as well as firm fixed effects to control for structural and time-invariant differences in stock returns at the firm level. We cluster standard errors at the firm level to control for serial correlation of the error terms.

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

[29] In the same fashion, we extract residuals from an AR(6) model from our monthly EnvP sentiment index.

[30] Scope 1 emissions are direct emissions from production, as opposed to scope 2 emissions which are indirect emissions from consumption of purchased electricity, heat or steam.

[31] We deflate all financial variables by annual GDP collected from the database of the St. Louis Fed.

Table 3.5: Estimation results - EnvP and Excess Stock Returns

	(1) Exc. ret.	(2) Exc. ret.	(3) Exc. ret.	(4) Exc. ret.	(5) Exc. ret.	(6) Exc. ret.
EnvP	0.0022***	0.0027***	0.0051***	0.0027***	0.0014***	0.0019***
	(0.0003)	(0.0003)	(0.0011)	(0.0003)	(0.0003)	(0.0003)
EnvP × CO_2 Emissions	-0.0003*	-0.0004***				-0.0004***
	(0.0001)	(0.0001)				(0.0001)
Quartile of emissions=2 × EnvP			-0.0016			
			(0.0012)			
Quartile of emissions=3 × EnvP			-0.0030***			
			(0.0012)			
Quartile of emissions=4 × EnvP			-0.0030***			
			(0.0011)			
EnvP × CO_2 Emission Intensity				-0.0003**		
				(0.0001)		
EnvP × Pre-sample CO_2 emissions					-0.0006***	
					(0.0002)	
EnvP net sentiment						-0.0043***
						(0.0005)
EnvP net sentiment × CO_2 Emissions						0.0001
						(0.0002)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year Trend	Yes	Yes	Yes	Yes	Yes	Yes
Firm controls	No	Yes	Yes	Yes	Yes	Yes
Risk factors	Yes	Yes	Yes	Yes	Yes	Yes
Observations	91,531	49,143	49,143	49,143	34,079	49,143
Firms	1400	614	614	614	262	614
R^2	0.54	0.72	0.72	0.72	0.77	0.72

The table presents results of an OLS regression. Standard errors are clustered at the firm level. The dependent variable corresponds to excess returns as continuously compounded monthly returns in excess of the safe rate. Emission measures refer to scope 1 (fixed mean) CO_2 emissions or emission intensity at the firm level. Firm controls include size as log(market capitalization), profitability as log(return on assets), leverage as log(total debt over total equity) and log(dividends per share). Risk factors include the market risk factor, SMB, HML, RMW and CMA. The EnvP innovations and EnvP sentiment innovations correspond to the residuals from an AR(7) and AR(6) process, respectively, and are standardized to a mean of zero and unit standard deviation. The emission measures are standardized in the same way. Standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.5 presents the results of our baseline estimation. Column (1) includes firm fixed effects, Fama-French risk factors and industry-year trends. We add firm controls in columns (2)-(7), which reduces the sample to about 600 firms. Columns (3)-(5) test different policy exposure measures such as quartiles of (fixed mean) CO_2 emissions, CO_2 emission intensities and pre-sample CO_2 emissions. In column (6) we add controls for innovations in our sentiment index. Across all specifications, we find that our coefficient of interest, i.e. the interaction term between our EnvP index and CO_2 emissions, has the expected sign and is highly significant at the 1 percent level. There is a negative association between our EnvP index and the stock returns of high-emission firms with greater exposure to environmental policy. Quantitatively, firms with CO_2 emissions one standard deviation above the sample mean experience a differential drop in excess returns of around 4 basis points following a one standard deviation (sd) EnvP innovation. For a firm with average exposure[32], however, news on environmental regulations tend to be positively associated with stock returns, as indicated by the positive coefficient on EnvP.[33] In column (3), we find that there are highly significant differences between the least polluting quartile and the two most polluting quartiles, with firms in the two highest polluting quartiles experiencing the largest relative drop in excess returns when EnvP rises relative to the least polluting quartile of firms. Moreover, in columns (4)-(5), we show that this effect is robust to using emission intensity or pre-sample emissions as a policy exposure measure. Finally, column (6) shows that the negative coefficient on the interaction term between our EnvP index and CO_2 emissions is robust to controlling for sentiment about environmental news, confirming again that our EnvP index has a negative significant effect on the stock returns of firms with greater exposure to environmental policy regardless of the monthly sentiment of EnvP news. For a firm with average exposure, however, an increase in positively toned news about environmental policy is associated with a drop in excess returns, as indicated by the negative coefficient on the EnvP net sentiment variable. As far as positive sentiment reflects a strengthening of environmental policy, this is aligned with the intuition that more stringent environmental regulations may be costly for most (average) firms. However, we do not find evidence of a significant difference for firms most exposed to environmental policy.

3.4.2 Environmental Policy and Aggregate Clean Investments

Having looked at firm-level estimations, we now consider the association between our index and aggregate investments in the clean energy sector at the macro level. By contrast to within-firms decisions, this may capture additional channels (e.g. entry and exit) of the dynamic relationship between investments and the salience of environmental policy.

Aggregate cleantech venture capital deals

We extract data on the monthly number of venture capital deals in renewable energy (which includes solar, wind, hydro and geothermal) over the January 1998 - March 2019 period from the i3 Cleantech Group database.[34] Since we focus on renewable energy, we use our index on ‘renewable energy policy’, as this is likely the most relevant for investors. Figure 3.9 plots our EnvP-RE index together with the aggregate monthly number of VC deals in renewable energy. Both series share a similar trajectory since the beginning of the 2000s, only diverging during the global financial crisis in 2008-2009 and over the 2015-2017 period.

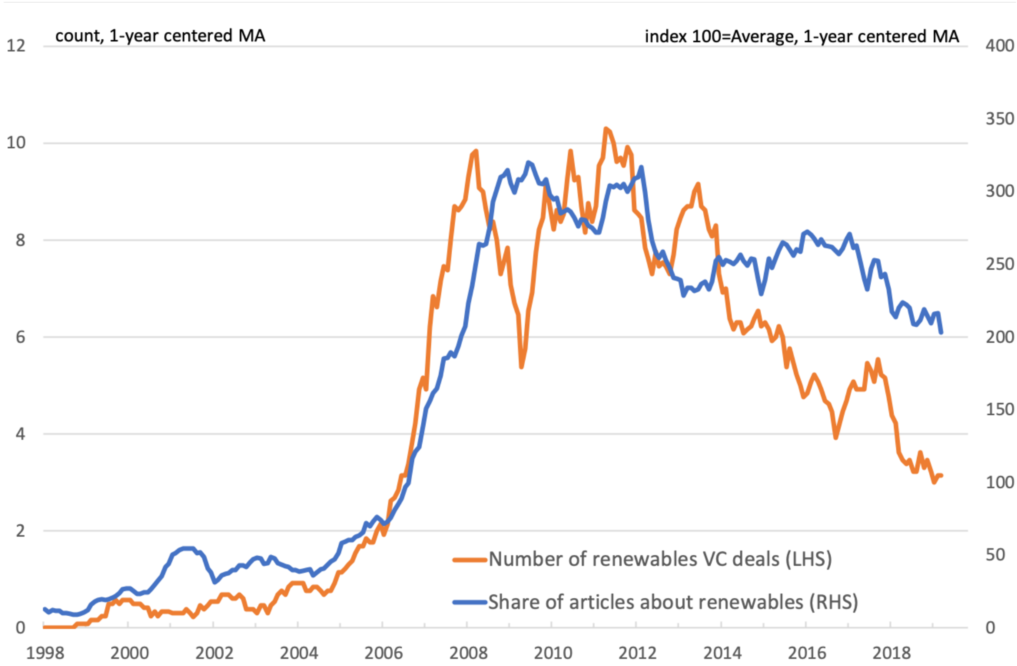
Our baseline VAR specification includes the following controls, all at monthly frequency: 1) oil prices as the West Texas Intermediate crude oil spot price from the St-Louis FED, 2) market risk captured by the Federal Reserve effective funds rate from the Board of Governors of the

[32]In the estimation, we standardize the CO_2 emissions variable, with mean zero and unit sd.

[33]We investigate this further by controlling for news sentiment in column (6).

[34]This database provides information on early-stage financing of 11,620 US cleantech startups (seed, series A, series B and growth equity) tracked over time by the Cleantech Group.

Figure 3.9: Evolution of number of renewable energy VC deals and EnvP-RE news index, monthly



Federal Reserve System, 3) aggregate economic activity using Markit's U.S. monthly real GDP index[35] and 4) a linear time trend. We include three lags of all variables, based on lag selection criteria. Table C.3 in Appendix C provides summary statistics of the variables in our sample.

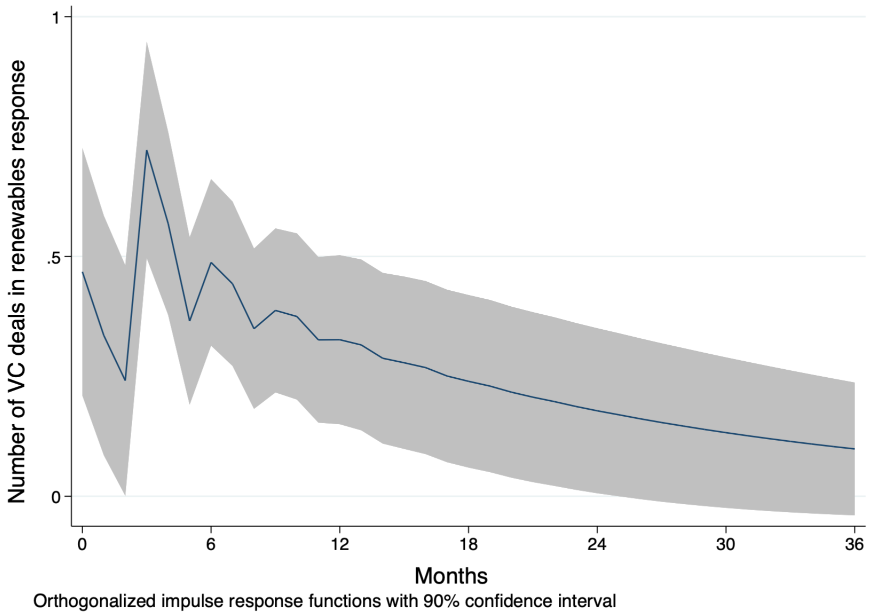
We conduct standard unit roots tests and use the monthly first difference of the following series, the log of oil prices, the log of GDP and the Federal funds rate, because these are not stationary in levels. As we can reject the presence of a unit root for the number of VC deals and the EnvP-RE news index using the Phillips–Perron test, we keep these two variables in levels in our preferred specification.[36] In order to recover orthogonal shocks we use the following Cholesky ordering: EnvP-RE news index, $\Delta \ln(\text{oil price})$, $\Delta \ln(GDP)$, Δ effective Fed funds rate, VC deals in renewable energy.

Figure 3.10 displays the model-implied impulse response function of the number of VC deals in renewable energy to a shock in our news-based EnvP-RE index. We see that a one standard deviation increase in our index is associated with about 0.6 more VC deals in the medium term. While this effect is moderate in size, it still represents a nearly 15% increase in the average monthly number of VC deals in renewable energy (i.e. 4.2 between January 1998 and March 2019). We show that this positive relationship between the EnvP-RE news index and VC

[35]For our robustness analysis below using Californian data, we use the Federal Reserve Bank of Philadelphia's coincident economic indicator, which includes non-farm payroll employment, the unemployment rate, average hours worked in manufacturing and wages and salaries.

[36]We expect VC deals which take several months to close to be more strongly correlated with the level of environmental policy in the media rather than its monthly evolution.

Figure 3.10: Estimated effect of a shock in EnvP-RE news on the number of renewable energy venture capital deals (IRF)



investments in renewable energy is robust to varying specifications on Figure C.6 in Appendix C.

Aggregate clean energy stocks

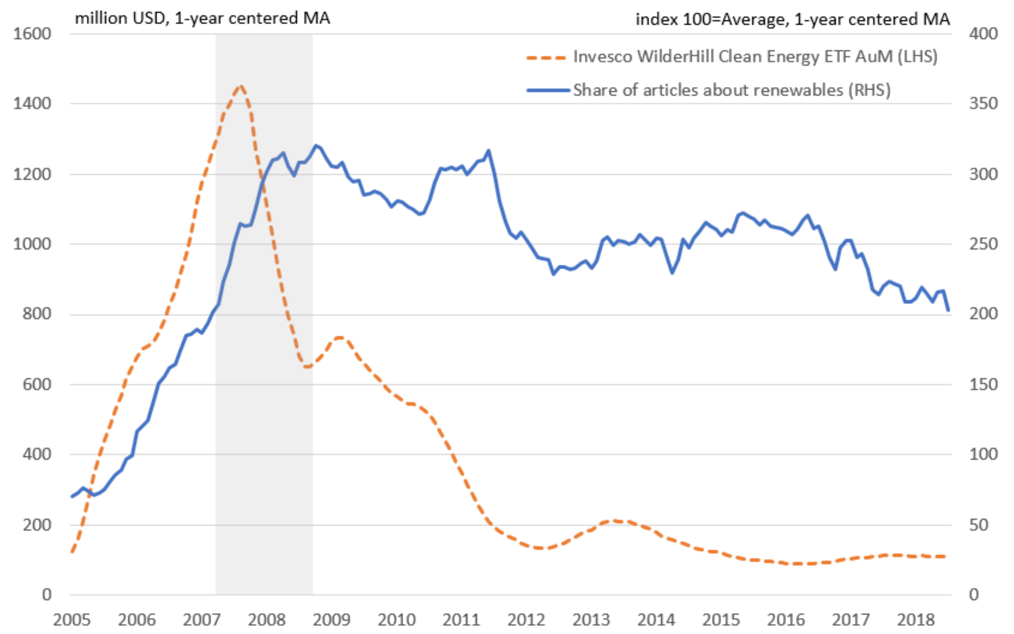
Next, we investigate the dynamic relationship between our news-based index and aggregate clean energy stocks. Specifically, we examine how the assets under management (AuM) of the Invesco WilderHill Clean Energy Exchange Traded Fund (PBW-ETF), tracking the portfolio of 52 US renewable energy companies, is associated with our index.[37] Considered as the main benchmark clean-energy index, the PBW-ETF is widely used in the energy economics literature (Kyritsis and Serletis, 2019; Sadorsky, 2012; Kumar et al., 2012). We extract this series from Datastream. Again, given the focus on renewable energy, we use our specific EnvP-RE index to measure news on renewable energy policy.

Figure 3.11 plots the monthly sub-index on renewable energy policy together with the assets under management of the PBW-ETF for the period of March 2005 to March 2019. The figure shows that the co-movement patterns of the series vary substantially over time, with a high

[37] The reason we are focusing on an ETF are twofold. First, one cannot directly invest in a market index. Second, it allows us to analyze the investment behaviour of less sophisticated investors who are more likely to learn something new from journal articles because retail investors are the main participants in the ETF market. A caveat of using the assets under management of an ETF is that they may be driven by fund flows or by changes in the value of the underlying assets. Therefore, it is a measure of demand for renewable energy stocks likely suffering from some measurement error.

co-movement before the Global Financial Crisis (GFC) but less so during and after it.[38]

Figure 3.11: Evolution of PBW-ETF and EnvP-RE news index, monthly. The shaded area corresponds to the recession following the Global Financial Crisis from December 2007 to June 2009.



Our baseline VAR specification includes the monthly assets under management of the PBW-ETF, EnvP-RE news and other controls as in (Kyritsis and Serletis, 2019; Sadorsky, 2012; Kumar et al., 2012), namely: 1) oil prices, as the US West Texas Intermediate crude oil spot price, 2) technology stocks, using the NYSE Arca Technology Index (PSE), and 3) market risk captured by the Federal Reserve effective funds rate. We exclude the recession associated with the GFC (December 2007 - June 2009) from the analysis.

As before, we run a series of unit root tests (Augmented Dickey–Fuller, Phillips–Perron and Kwiatkowski–Phillips–Schmidt–Shin tests). Accordingly, we use the PSE and oil prices in monthly log differences, PBW-ETF AuM and the effective Fed funds rate in monthly differences and, although tests are less conclusive in this case, we use the EnvP-RE index in monthly log differences in our baseline – so we consider changes in the growth rate of our EnvP index. We include one lag for all variables as suggested by standard tests and recover orthogonal shocks by imposing the following Cholesky ordering: $\Delta\ln$(EnvP-RE index), $\Delta\ln$(oil price),

[38]The correlation of the annual centered moving average of PBW-ETF AuM and our news index between 2005 and 2007 is very high at 0.9. During the recession caused by the Global Financial Crisis (GFC), officially dated by the NBER from December 2007 - June 2009, PBW-ETF AuM take a dip, while policy news about renewable energy remain at elevated levels amid the announcement of the Green New Deal and Obama's era of strong support for renewable energy. The correlation of PBW-ETF AuM and our news index during this time period is at -0.9. The post-GFC period is marked by a much lower co-movement of the PBW-ETF AuM with our EnvP-RE policy news index at 0.7 (0.6).

ΔFederal Reserve effective funds rate, $\Delta \ln(\text{PSE})$, $\Delta \ln(\text{ETF-PBW})$.[39] Summary statistics are provided in Table C.6 in Appendix C.

Figure 3.12: Estimated effect of a shock to the growth rate of the EnvP-RE news index on the change in assets under management of the PBW exchange traded fund (IRF)

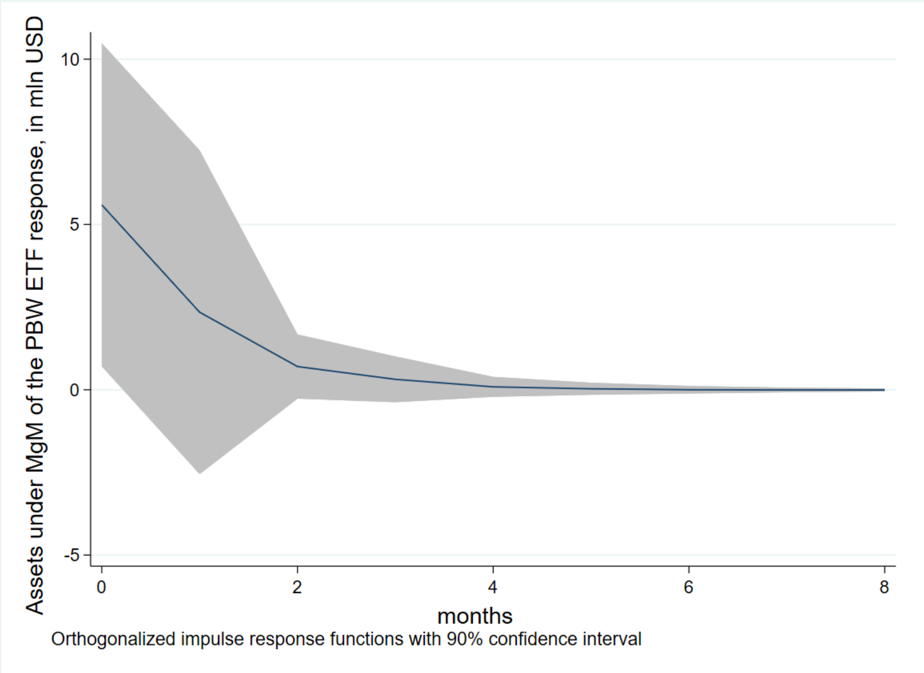


Figure 3.12 shows that a one-standard deviation shock to the growth rate of our EnvP-RE index is associated with an additional increase of 5 million USD in assets under management of the PBW-ETF. While this effect seems rather small, it still represents a 125% increase in the average monthly change in AuM of the PBW ETF (i.e. about 4 mln USD between April 2005 and March 2019). The result is broadly in line with the previous literature which finds a quantitatively small dynamic relationship between investor sentiment in renewable energy, as measured by the Google Trends Search Volume Indices and Tweets, and clean energy stock returns (Reboredo and Ugolini, 2018; Song et al., 2019).[40] Figure C.7 in Appendix C shows that our results are robust to various other specifications.

[39] Akaike information criterion (AIC), Final Prediction Error (FPE) and Hannan-Quin information criterion (HQIC)

[40] Investors in renewable energy markets may instead be more responsive to factors that move technology stocks than to environmental regulation. Sadorsky (2012), for instance, point out that renewable energy companies tend to behave similarly to high-tech companies because their success hinges on very specific technologies. Consistent with this hypothesis, we find that the PSE and PBW-ETF AuM have a positive association of about double the size of the link between EnvP-RE and PBW-ETF. Moreover, the link between oil prices and PBW-ETF AuM is about the same size as the one between EnvP-RE and PBW-ETF, in line with the notion that rising oil prices trigger a substitution towards renewable energy technologies (Kumar et al., 2012; Sadorsky, 2012).

3.5 Conclusion

Quantifying fine-grained information on environmental policy over several decades has often proven difficult. We apply text-mining techniques to newspapers archives to develop the EnvP index, a novel index measuring the salience of US environmental policy over the 1981-2019 period. The index captures the evolution of the relative share of news articles discussing environmental and climate regulations over the last four decades. We perform several reality checks showing that our index captures trends and peaks in the historical evolution of US environmental policy and co-moves with the stringency of the regulatory framework in a meaningful way. From a methodological point of view, our analysis showcases how newspaper archives combined with machine learning algorithms for text classifications can be exploited to retrieve a vast and diverse amount of information on environmental policy. In addition, the algorithms provide a much improved methodology, compared to simpler information retrieval using keywords. We illustrate how the index can be further exploited to build many additional indicators, providing information on sentiment (i.e the tone of articles) and on sub-topics such as ‘renewable energy policy’ and ‘international climate negotiations’.

We further look at how our EnvP index relates to financial investments in clean markets. Our results provide a range of empirical evidence corroborating that our news-based measure of environmental policy has a meaningful association with clean investments as proxied by venture capital financing and stock returns – both in firm-level panel estimations and VAR models. More specifically, a doubling of environmental policy news is associated with a 26% increase in the likelihood than an average cleantech startup receives funding. Conversely, a one-sd increase in (an innovation in) our EnvP index is associated with a loss of about 4 basis points in excess returns for the most polluting firms. Furthermore, we find in VAR models that a shock in our sub-index on renewable energy policy is associated with an increase in the number of clean energy deals at the macro level and an increase in the assets under management of the main clean energy exchange-traded fund. The persistence of our findings across many specifications serves as validation test that our EnvP index captures a meaningful association with clean investments.

An immediate implication of our results is that our EnvP index can help assist the policy and finance community to better understand which type of policy signals (e.g. design, compensation schemes) may be most effective in affecting investor beliefs. News on environmental policy appear to contain a lot of relevant information for investors and policymakers could leverage their capacity to coordinate investor beliefs by communicating about their environmental policy agenda in a clear and credible manner, akin to central banks coordinating inflation expectations through forward guidance. In this respect, our study offers many avenues for future research. We hope in particular that our index can help researchers to progress towards quantifying causal impacts of specific features of environmental regulations, for instance by combining our index with event studies or quasi-natural experiments. In a more macro setting, our index may provide an improved quantification of transition risks in the context of climate change and the low-carbon transition.

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

This chapter is written in collaboration with Joëlle Noailly and Laura Nowzohour.[1]

Abstract:

We use text-as-data methods relying on supervised machine learning algorithms applied to newspapers to construct a new index of policy uncertainty in US environmental and climate regulations over the 1990-2019 period. We find that policy uncertainty is pervasive in the history of US environmental policy and tends to rise around salient events and election cycles. We further examine how our environmental policy uncertainty index – the EnvPU index – relates to investments in the low-carbon economy, which are by nature more dependent on policy support. In firm-level estimations, we find that our index is associated with a reduced probability for cleantech startups to receive venture capital (VC) funding, in particular for clean energy startups characterized by capital-intensive investments that are difficult to reverse. In financial markets, a rise in our EnvPU index is associated with higher stock volatility for firms with above-average green revenue shares. At the macro level, shocks in our index lead to declines in the number of cleantech VC deals and higher volatility of the main clean energy exchange-traded fund. Overall, our results are consistent with the notion that policy uncertainty delays investments for the low-carbon economy, thereby representing a key bottleneck in addressing the urgency of climate action.

[1] The three co-authors contributed equally to the core of the paper: the research question, framing, writing and labelling of newspaper articles. Hired research assistants also contributed to the labelling. The doctoral candidate and L.N implemented the SVM and organized the human audit. J.N is the main contributor to the introduction. L.N is the main contributor to the stock market analysis (in Section 4.4). The doctoral candidate is the main contributor to the venture capital analysis (in Section 4.4) and the political cycles validation exercise (Sections 4.3.3 and 4.3.4).

4.1 Introduction

In recent decades environmental regulations have alternated between periods of increasing stringency and sudden rollbacks. Over the course of its mandate, the Trump administration not only announced its withdrawal from the Paris climate agreement but also conducted a comprehensive review of many federal environmental regulations. These changes created a surge in uncertainty about the state of future environmental regulations as it became unclear how and when they would be implemented, how many policies would be dismantled and whether these rollbacks would be legally challenged in the future. Other countries conducted similar rollbacks of their environmental and climate policies. France revoked its fuel tax after the Yellow Vests protests and Australia and China recently backpedalled on coal regulations. While policy revisions in response to new information are inevitable and desirable, the politics of environmental policy are particularly volatile. Policymakers often have to balance competing (long-term) environmental objectives with (short-term) economic and electoral priorities. As environmental regulations typically face a lot of opposition in the form of lobbying, protests, legislative battles or legal challenges, even advanced and promising policy proposals may, in the end, fail unexpectedly. Such abrupt and unpredictable changes generate substantial uncertainty, making it difficult to anticipate how the regulatory framework will unfold in the future.

Given the urgency of climate action, environmental policy uncertainty is a major obstacle to a rapid and timely transition to a low-carbon economy. Firms typically want to know which policy assumptions will be valid for their projects over the next decades and, when faced with a policy framework characterised by high uncertainty, may prefer to postpone or withdraw their investment. According to a recent survey from the European Investment Bank, 43 percent of European firms and 22 percent of US firms cite ‘uncertainty about regulation’ as an important barrier to undertaking climate-related investment. European firms rank policy uncertainty as the most important obstacle, while for US firms policy uncertainty comes just after ‘investment costs’ but above ’availability of finance’ (European Investment Bank, 2021). Understanding how policy uncertainty affects and delays clean investments is essential to provide better guidance on the significance of the timing and credibility of policy action (Goulder, 2020).

The objective of this study is to examine the relationship between environmental policy uncertainty and investments in the low-carbon economy. We construct a news-based index of environmental policy uncertainty (EnvPU) by using supervised machine learning algorithms on articles about environmental policy extracted from the archive of ten US newspapers over the 1990-2019 period. More precisely, our EnvPU index represents the share of regulatory uncertainty news among environmental policy articles in a given month. Relying on newspapers is a significant advantage of this study as they provide a broad set of textual information on environmental policy at a high frequency (daily, weekly, monthly) over long periods. Because newspapers allow us to continuously track environmental policy over time, we can compare the intensity of policy uncertainty over various periods. This is particularly important for a transient concept like uncertainty. In addition, journalists' writing directly affects how readers the uncertainty around given topics. As a result, using newspaper articles is a great way through which to tap into investors' state of mind. Our main contribution is to propose a new methodology based on machine learning algorithms for text classification, enabling us to retrieve meaningful information from text accurately.

The historical evolution of US environmental policy uncertainty as reflected by our index shows several prominent peaks. Our EnvPU index spikes at the end of 1995 when a disagreement over cuts in environmental regulations led to a government shutdown for several weeks. Environmental policy uncertainty then plummeted in the mid-2000s. The end of the decade was then punctuated by several bursts of policy uncertainty related, among other things, to legislative hurdles around the climate bill aiming to introduce a national cap-and-trade system, the failure of the COP15 in Copenhagen, legal challenges between Texas and the EPA, and new rules regarding offshore drilling after the Deepwater Horizon oil spill. The latest spikes in environmental policy uncertainty took place as expected during the Trump presidency after 2016.

We evaluate our index in several ways. First, we compare the performance of our index to an alternative index based on a simple keyword approach and demonstrate the added-value of machine learning algorithms in terms of precision and recall. Second, we assess the performance of our algorithm by benchmarking it to a human audit study which confirms the validity of our index. Third, we provide evidence that our EnvPU index correlates with periods of objective policy uncertainty, i.e., significant shifts in pro-environmental votes in Congress along the US electoral cycle. Finally, we verify that our index is not affected by political slant in newspaper coverage.

Conceptually, we expect that increases in environmental policy uncertainty lead to lower investments in clean technologies. However, measuring the impact of environmental policy uncertainty on investments could raise endogeneity concerns. Separating environmental policy uncertainty from other types of (omitted) uncertainty such as technological or climate change uncertainty, which likely correlate with investments, is indeed challenging. Still, we expect that the correlation between these omitted types of uncertainty and our EnvPU is small – motivated by the fact that we expressly excluded non-policy-related uncertainty in our text classification task. Moreover, potential concerns about reverse causality are mitigated by the fact that variations in our index are, at least partly, driven by (exogenous) political elections. Our firm-level identification strategy relies on differentiating firms according to their level of exposure to environmental policy uncertainty. We argue that investments in the firms most exposed to environmental policy uncertainty should exhibit a stronger co-movement with the EnvPU index. Indeed, we find that our EnvPU index is negatively associated with VC investments in cleantech startups. The adverse effect of policy uncertainty is particularly strong for the sub-group of clean energy startups, characterized by capital-intensive investments that are difficult to reverse. Finally, changes in the EnvPU index have a negligible impact on VC funding of non-cleantech startups. In financial markets, a rise in our EnvPU index is associated with higher stock volatility for firms with above-average green revenue shares. Then in VAR settings, we find similarly that a shock in EnvPU leads to reduced clean energy VC deals at the aggregate level and higher volatility of the main clean energy exchange-traded fund.

Our work relates to several strands of literature. First, our paper links to the large body of theoretical work looking at the effects of uncertainty on various aspects of economic activity. A well-established result is that firms faced with uncertainty tend to adopt a wait-and-see behavior and postpone their investments (e.g., Bernanke, 1983; Mcdonald and Siegel, 1986; Dixit et al., 1994; Bloom, 2009; Bloom et al., 2018). Uncertainty shocks lead firms to reduce or delay their investments in particular when these are costly to reverse, affecting productivity growth and output. Theoretical models in environmental economics mainly consider uncertainty from environmental and climate shocks, which lead to increases in precautionary savings, capital adjustments (e.g., away from capital prone to climate shocks) and possibly lower productivity, growth and welfare (Bakkensen and Barrage, 2021; Cai and Lontzek, 2018).

Theoretical work looking specifically at policy uncertainty finds in a similar fashion that firms tend to postpone their investment when the regulatory framework becomes uncertain, with the pent up investment generating a subsequent boom (Stokey, 2016). In financial markets, policy uncertainty depresses investments by raising risk premia (Pástor and Veronesi, 2012). Only a handful of theoretical papers has extended the analysis to uncertainty about environmental policies. In an early work, Viscusi (1983) shows that if investments are irreversible, uncertainty about environmental policies causes firms to invest in fewer pollution controls. Bretschger and Soretz (2018) find that policy uncertainty leads to increases to precautionary savings and diversification of portfolio away from risky assets that could potentially become stranded in the future. Other models have used real option methods to analyze the impact of uncertainty in carbon pricing policies on investments in coal- and gas-fired power plants, finding that policy uncertainty creates a risk premium for power investment (Blyth et al., 2007; Fuss et al., 2008).

Our work also relates to the extensive empirical literature on the impact of policy uncertainty on investment (Julio and Yook, 2012; Durnev, 2010). A major empirical contribution derives from the work of Baker et al. (2016) who construct a continuous high-frequency index of economic policy uncertainty by counting the frequency of articles containing relevant keywords in newspapers. Increases in their news-based index of economic policy uncertainty (EPU) are associated with lower investments and increases in stock volatility for firms most exposed to policy uncertainty. At the macro level, their index is associated with reductions in industrial production and employment. In recent years, a large body of empirical work has found that a rise in the EPU index is associated with declines in firm-level (irreversible) investments (Guien and Ion, 2016), mergers and acquisitions (Bonaime et al., 2018), and venture capital funding (Tian and Ye, 2017). In financial markets, other studies find that the EPU index commands a risk premium (Pástor and Veronesi, 2013) and helps to forecast market returns (Brogaard and

Detzel, 2015).

In environmental economics, empirical studies tend to rely on event studies to examine how environmental policy uncertainty affects investments. Lemoine (2017) exploits the unexpected collapse of the cap-and-trade climate bill in April 2010 to show that it led to an increase in coal prices and inventories. Sen and von Schickfus (2020) find that uncertainty about the implementation of a compensation mechanism for a carbon fee for energy companies led to an abrupt devaluation of companies holding fossil fuel assets. Dorsey (2019) exploits a legal challenge to the Clean Air Interstate Rule and finds that plants with a lower probability of being regulated reduced pollution by 13 percent less and compliance costs overall increased by $124 million due to efficient investments being delayed. Besides event studies, other papers use volatility in annual energy R&D expenditures (Kalamova et al., 2013) or survey-based perceptions of investors (Johnstone et al., 2010) to measure environmental policy uncertainty but face the challenges of low-frequency data and identification issues in empirical work.

Finally, we relate to the rapidly growing literature using text-as-data in economics (Gentzkow et al., 2017; Dugoua et al., 2021), particularly to several applications on measuring the salience of environmental policy in newspapers (Noailly et al., 2021) and identifying firm-level climate risks (Sautner et al., 2020; Kölbel et al., 2020) based on firms' 10k filings and earnings calls.[2] Closer to our work on policy uncertainty, Tobback et al. (2018) reproduce the EPU index of Baker et al. (2016) using a supervised algorithm just as we do. Basaglia et al. (2021) construct a climate policy uncertainty index over 1990-2018 by counting the frequency of articles combining keywords on uncertainty and climate policy. These authors investigate the association between their climate policy uncertainty index and firm-level outcomes for the US and find that their index is linked to larger stock price volatility and lower share prices for firms in emission-intensive industries. They also find that climate policy uncertainty is negatively associated with R&D and annual employment. Compared to their approach, we provide a more refined methodology to identify articles about environmental policy uncertainty based on machine learning techniques. We also use a different conceptual framework by arguing that it is the greenest, not the most pollution-intensive firms, which are likely to be most affected by climate policy uncertainty.

The paper is organized as follows. Section 4.2 describes our methodology and text-mining algorithm and describes our index of environmental policy uncertainty. Section 4.3 provides a set of validity checks. Section 4.4 explores how our index relates to low-carbon investments and Section 4.5 concludes.

[2] Climate risks or more specifically 'transition risks' refer to the risks of a significant strengthening of environmental policy in the future. This is different than our environmental policy uncertainty measure, which captures the inability to predict how future regulations will look like, let alone whether a strengthening or relaxing of regulations will take place.

4.2 Measuring Environmental Policy Uncertainty

In this section, we present the various steps towards constructing our news-based index of environmental policy uncertainty using automated methods. In contrast to existing studies (Baker et al., 2016; Basaglia et al., 2021), we rely on supervised machine learning instead of using a subjective set of keywords to identify policy uncertainty. We show that this approach significantly improves the quality of our classification. We are able to capture more fine-grained trends and peaks in the history of US environmental policy uncertainty news. Comparing our index to two indices created using a keyword-based approach and a comprehensive human audit exercise further underlines its validity and added-value. Finally, we also compare our index to electoral cycles and fluctuations in pro-environmental votes.

4.2.1 Environmental policy news

Our dataset is a sample of 80,045 newspaper articles about US environmental policy, which we have identified in previous work (Noailly et al., 2021). Our previous *environmental policy* classification exercise started from 15 million news articles extracted from the archives of ten leading US newspapers over the period from January 1981 to March 2019, via automated access through Factiva.[3] This first classification also relied on supervised machine learning algorithms. We show in Noailly et al. (2021) that this approach yields excellent performance metrics and provide a large set of validity checks.

Figure 4.1 shows our index of environmental policy, which corresponds to the monthly share of all news articles that are about environmental policy, normalized to an average value of 100 over the 1981-2019 period. The index correctly captures salient events in the history of US environmental policy, such as the Green New Deal during Obama's Presidency, major UNFCCC climate change conferences, or Trump's announcement of withdrawal from the Paris agreement in June 2017. In the remainder of our analysis, we use the EnvP index to control for the level (first-moment) of environmental policy, whereas our newly constructed environmental policy uncertainty index reflects how volatile and unpredictable (second-moment) the regulatory framework is.

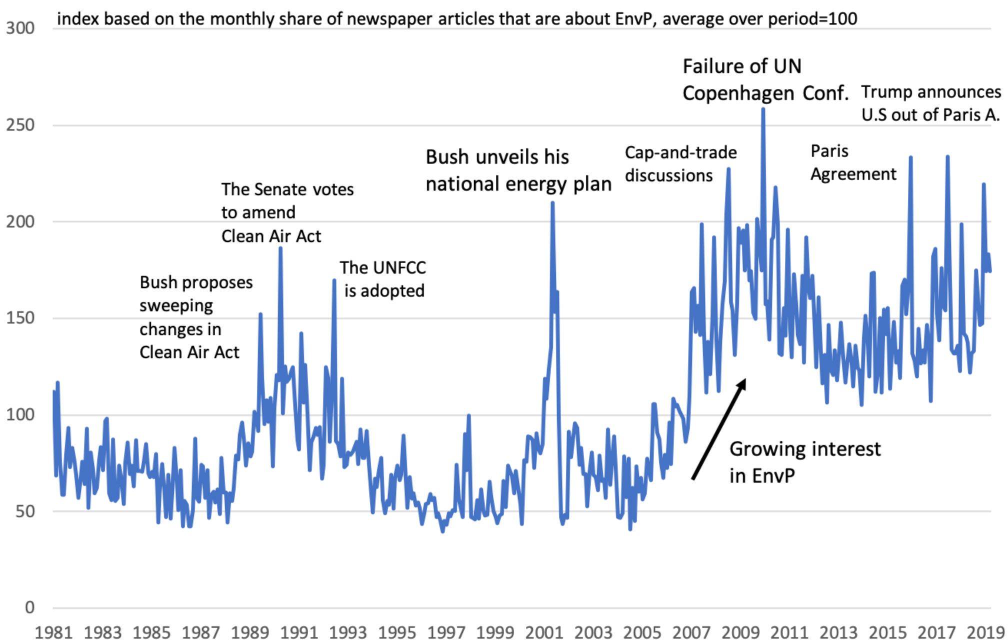
4.2.2 Developing an index of environmental policy uncertainty

Training set, text pre-processing and classification

Within the set of 80,045 EnvP articles described in Section 4.2.1, we aim to identify the subset of articles about environmental policy *uncertainty*. First, we generate a training set of articles that reflects what constitutes a relevant article about environmental policy uncertainty and what does not. While reading a large set of articles, we develop a detailed codebook to guide

[3]The list of newspapers includes *New York Times, Washington Post, Wall Street Journal, Houston Chronicle, Dallas Morning News, San Francisco Chronicle, Boston Herald, Tampa Bay Times, San Jose Mercury News and San Diego Union Tribune*. Appendix D provides additional statistics on the set of newspapers.

Figure 4.1: EnvP - An index of environmental policy



the classification of EnvPU articles.

We only label articles as relevant if they pertain to *current* policy uncertainty. Typically, articles giving rise to perceptions of an inability to predict how future environmental policy will unfold fall into the following categories:

1. There is a policy shift or reversal of environmental regulations, creating volatility in the politics of environmental policy.[4] This type of uncertainty is inherent to elections and transitions in political cycles.[5]

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

   - political or business opposition which may block or slow down the legislative process and thereby the fate of the policy,
   - a lack of political will and supportive coalitions, for instance in the case of climate change negotiations,
   - legal challenges where the court's decision is still pending.

[4] We therefore do not label as uncertain articles that depict a clear and certain reduction in policy. We want to separate falls in stringency from uncertainty.

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

3. There is uncertainty about what the policy rules will exactly entail, when the policy will start, or whether the policy will be enforced.

4. Other types of policies (e.g. trade dispute on solar tariffs) may have uncertain impacts on clean markets.

Using this codebook, we manually label a random sample of 622 articles from our subgroup of 80,045 EnvP news articles.[6] Each article is labelled by at least two annotators. The result of the manual classification yields 204 articles labeled as relevant for environmental policy uncertainty, so about 30 percent of the training set.

In a second step, we apply standard text pre-processing techniques to our set of environmental policy news articles and convert each document into numerical vectors of unigram, bigram and trigram frequencies. We then construct a term-frequency inverse-document frequency (tf-idf) matrix, in which less weight is given to words that occur too often or too rarely.[7]

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

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

Using our best performing algorithm, we obtain an average precision of 56 percent and a recall of 70 percent, when predicting which of the 622 articles in our training set are about EnvPU.[8] How should we evaluate these metrics? First, with a precision of 56 percent, our algorithm performs significantly better than a random classifier (which would give a precision of about 30 percent given that less than one third of articles are classified as relevant). Second, because the task of distinguishing policy uncertainty articles within a first subset of environmental policy news (the two-step approach we adopt in this paper) is harder than that of distinguishing environmental policy uncertainty articles within a general set of news (the one-step approach commonly used in the literature), our precision metrics is lower than it could be.[9] Overall,

[6]The size of our training set might raise concerns that it is not be large enough to capture sufficient information on policy uncertainty. Nonetheless, we found that increasing the sample above 500 articles by increments did not significantly improve the performance of our classifying algorithm. Given that manually labeling articles is a very time-consuming and difficult task, we did not consider further extension of the training set.

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

[8]These performance metrics are an average of five different ten-fold cross validations using different random seeds. We use similar parametrization and cross-validation as in Noailly et al. (2021) In other words, more than half of the articles classified as EnvPU were also labelled as uncertain by the annotators (precision). Moreover, the EnvPU classifier successfully retrieve more than two thirds of all the articles labelled as uncertain (recall).

[9]It is significantly harder to precisely identify environmental policy *uncertainty* articles in a set of environmental policy news than it is to identify economic policy uncertainty articles in a broad pool of news covering various

Table 4.1: Top discriminating words for predicting our EnvPU index according to the trained SVM classifier.

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

identifying uncertainty is a complex and subjective task. Even with a codebook, two annotators would on average disagree on the label assigned to an article about 30 percent of the time. Hence, if humans cannot perfectly identify relevant articles, we cannot expect our algorithm to do so.

Descriptive statistics

To lend credence to our classifier, we provide a number of descriptive statistics such as the most important words used by the algorithm to make its classification choice, the articles which it most confidently classified as talking about EnvPU as well as the resulting index with labelled peaks.

Table 4.1 displays the most important text features used by the SVM classifier to predict whether an article falls into the ‘environmental policy uncertainty’ classification. It is reassuring to observe that this list of words encompass 1) environmental and climate issues (i.e. ‘emission’, ‘pipeline’, ‘drilling’, ’auto’, ’clean’, ’wind’), 2) policy-making (i.e. ‘epa’, ‘agency’, ‘rule’, ‘congress’, ‘administration’, ‘trump’, ‘clinton’) and 3) uncertainty terms (i.e. ‘court’, ‘review’, ‘cut’, ‘court’, ‘issue’, ‘fight’ or ‘delay’). These features are a sign that uncertainty arises either because of the presence of words associated with a specific uncertainty-inducing event (e.g., ‘trump’) or words that in themselves signal uncertainty (e.g., ‘delay’).

topics. Identifying a pattern about environmental policy uncertainty articles in the total volume of news would be much easier to grasp for the algorithm, but less accurate in terms of classifying based on policy uncertainty.

Table 4.2 reports excerpts of the five newspaper articles that were classified as the most likely to be about EnvPU. The first article titled *Trump officials deploy court tactic to reverse Obama rules* describes how President Trump was using both executives orders and court tactics to nullify numerous Obama-era green regulations. The second article *Court rebuffs Trump's effort to halt Obama methane rule*, reports on a federal court decision to prevent the EPA from suspending methane regulations. All of these articles describe either the rollback of environmental policies or legal battles over environmental regulations.

Table 4.2: Newspapers articles with the highest SVM-score

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

Next, we generate our index of US environmental policy uncertainty, which represents the monthly share of environmental policy uncertainty articles over all environmental policy articles, normalized to an average value of 100. Hence, an increase in our EnvPU index captures a meaningful increase in the prevalence of policy uncertainty amid ongoing debates on environmental policy. Our index is, therefore, unaffected by increasing media attention to environmental policy news. Scaling by the count of environmental policy articles rather than by the total volume of all news is an improvement over the standard approach as the latter would not allow us to differentiate increases in EnvPU due to more coverage to environmental policy (which by construction would raise the count of articles about policy uncertainty) from increases due to an actual rise in policy uncertainty.

Figure 4.2 plots the historical evolution of US environmental policy uncertainty over the 1990-2019 period. We find that our EnvPU index peaks at the end of 1995 when a disagreement over cuts in environmental regulations led to a government shutdown for several weeks. In this period, Republicans gained control of both houses of Congress for the first time since 1954 and attempted to push an anti-regulations agenda and block Federal agencies from imposing new rules on health, safety and the environment. Environmental policy uncertainty plummets in the mid-2000s, while the end of the decade is punctuated by several bursts in policy uncertainty.

Figure 4.3 zooms in on the 2009-2019 period to give a more fine-grained picture of what our index is able to capture. The first spike corresponds to the summer of 2010, during which our index was 40 percent above its average level. Uncertainty was then due to the introduction and eventual failure of a comprehensive climate bill sponsored by John Kerry, Joe Lieberman and, initially, Lindsay Graham. The second spike corresponds to the early months of Trump's presidency in 2017, where our index was 60 percent above its average level. At the time, there was great uncertainty about the extent to which Trump would dismantle environmental protection.

Figure 4.2: EnvPU - An index of environmental policy uncertainty 1990-2019

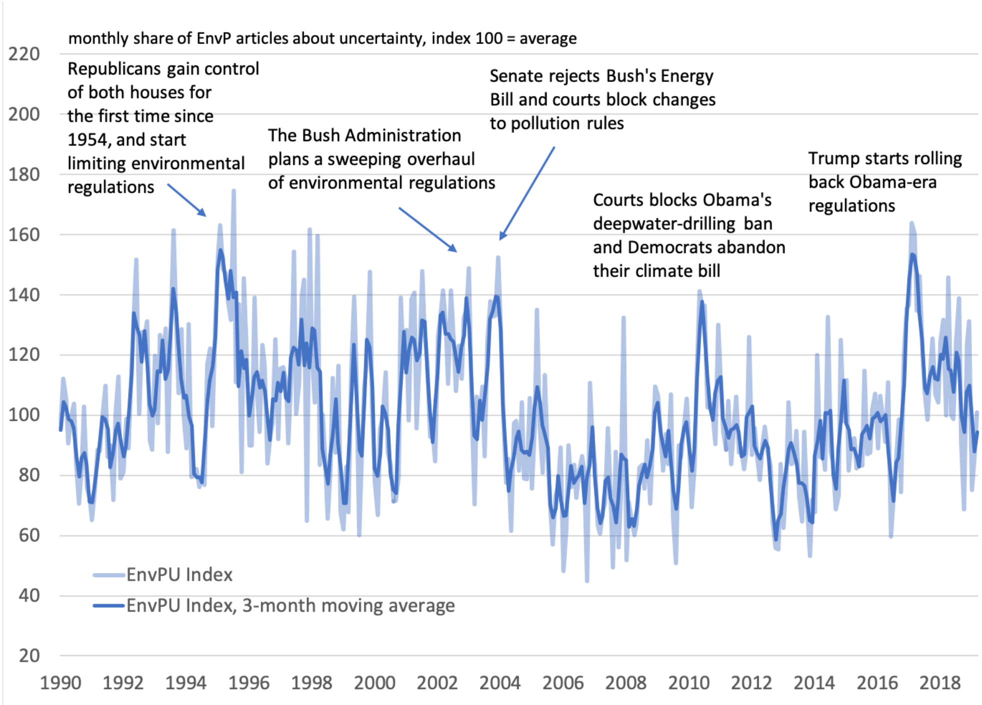
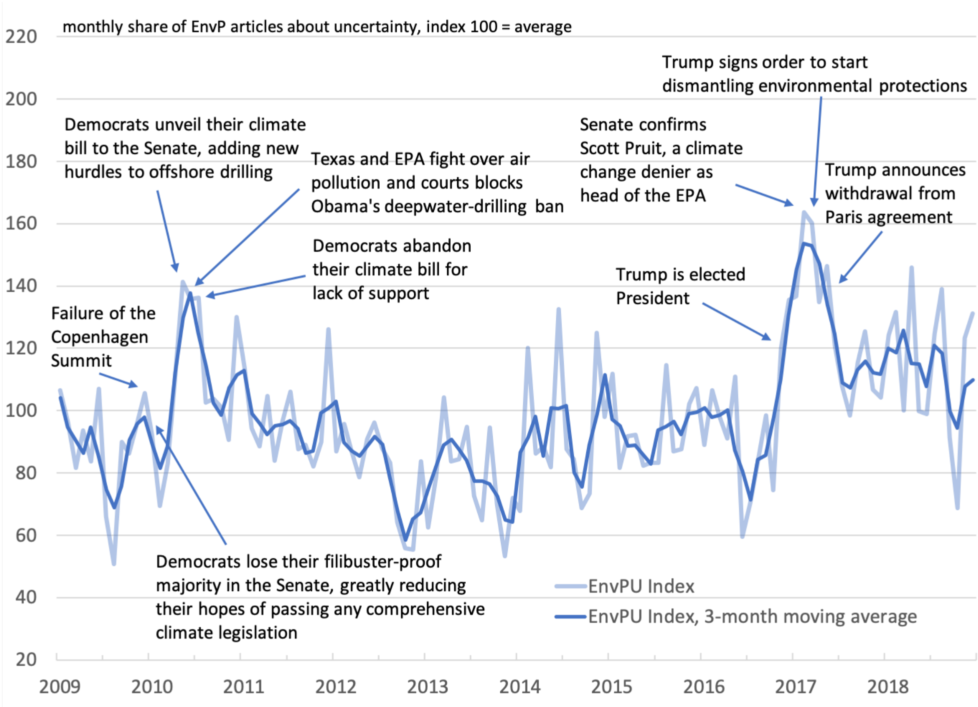


Figure 4.3: EnvPU - An index of environmental policy uncertainty 2009-2019



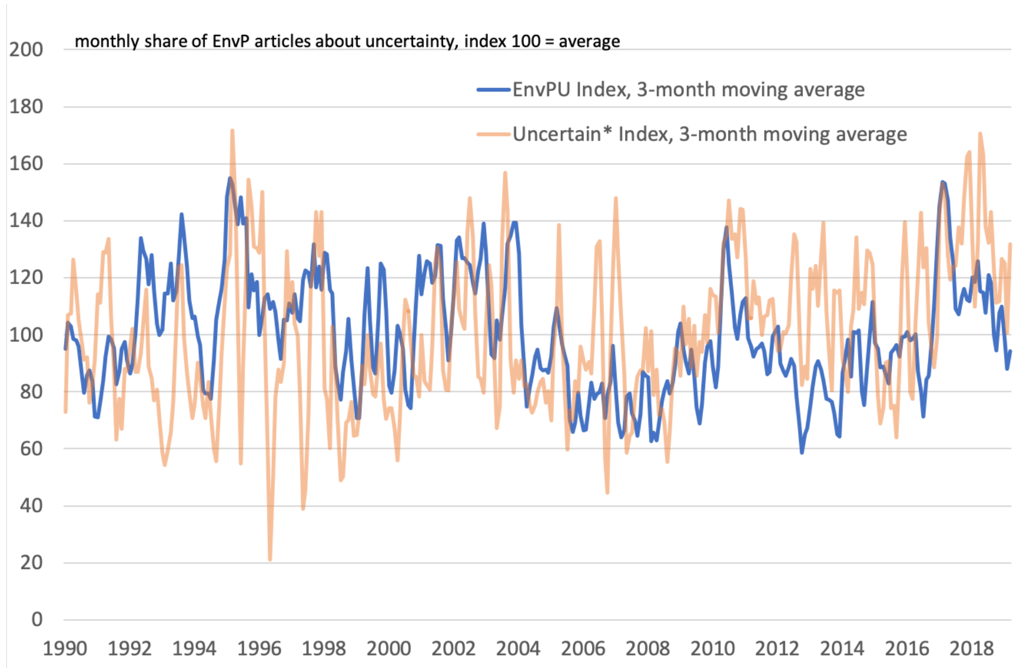
4.3 Evaluating our environmental policy uncertainty index

4.3.1 Comparison with a keyword-based approach

To benchmark our supervised machine learning index, we compare it with a 'naive' dictionary-approach which consists in searching for environmental policy articles including *uncertain** keywords as in Baker et al. (2016) and Basaglia et al. (2021). More specifically, we look for environmental policy articles containing one or more of the following four keywords *uncertain*, *uncertainly*, *uncertainty* or *uncertainties*.[10] We count the number of articles with *uncertain** keywords per month in the subset of 80,045 environmental policy news, scaled by the volume of environmental policy news. The index is standardized such that 100 represents the average over the 1990-2019 period.

Figure 4.4 plots the 3-months moving average of both the EnvPU and *uncertain** indices. The most striking difference is that the *uncertain** approach yields a much more volatile index oscillating around its average value of 100 and displaying fewer trends.

Figure 4.4: EnvPU versus *uncertain**, 3-month moving average.



Using our training set as a benchmark, we are able to compute the performance of the *uncer-*

[10]We also discuss in Appendix D a more elaborate classification, based on an extensive dictionary of keywords related to uncertainty. This method does not yield to better results than the two approaches discussed here. Tobback et al. (2018) compare – like we do – three similar methods to reproduce the EPU index from Baker et al. (2016): SVM, naive and a longer list of uncertainty modal words. Similar to us, they find that the latter approach is not significantly better than the first two.

*tain** approach, something other papers like Basaglia et al. (2021) cannot do without creating their own training sets. Using *uncertain** yields a precision of 49 percent and a recall of 8 percent. While the articles identified as EnvPU by this method will indeed be about environmental policy uncertainty nearly half of the time – lower than our precision of 56 percent – the recall is very low. Indeed, 92 percent of newspapers articles about uncertainty will be missed by this approach. This finding is aligned with Tobback et al. (2018) who compare various algorithms to reproduce the EPU index from Baker et al. (2016) and also find that the 'naive' approach suffers from a very low recall.[11]

Table 4.2 shows that most articles about environmental policy uncertainty do not use the term *uncertain** but instead use a wide lexicon of words related to uncertainty (e.g., President Trump [...] has reversed course on nearly two dozen environmental rules [...]). This comparison sheds new light on the downsides of the current standard keyword approach to capture policy uncertainty. While it is true that simple *uncertain** keywords can reliably identify some of the articles about policy uncertainty it will miss the vast majority of them, which begs the question of whether some topics are systematically omitted. Additionally, being entirely dependent on very few terms increases the volatility of the *uncertain** index, as Figure 4.4 illustrates. Our machine learning approach, because it is based on a wide array of features, is able to cast a much broader picture. As a result, it identifies many more articles about policy uncertainty than the *uncertain** approach, without weighing on its precision.

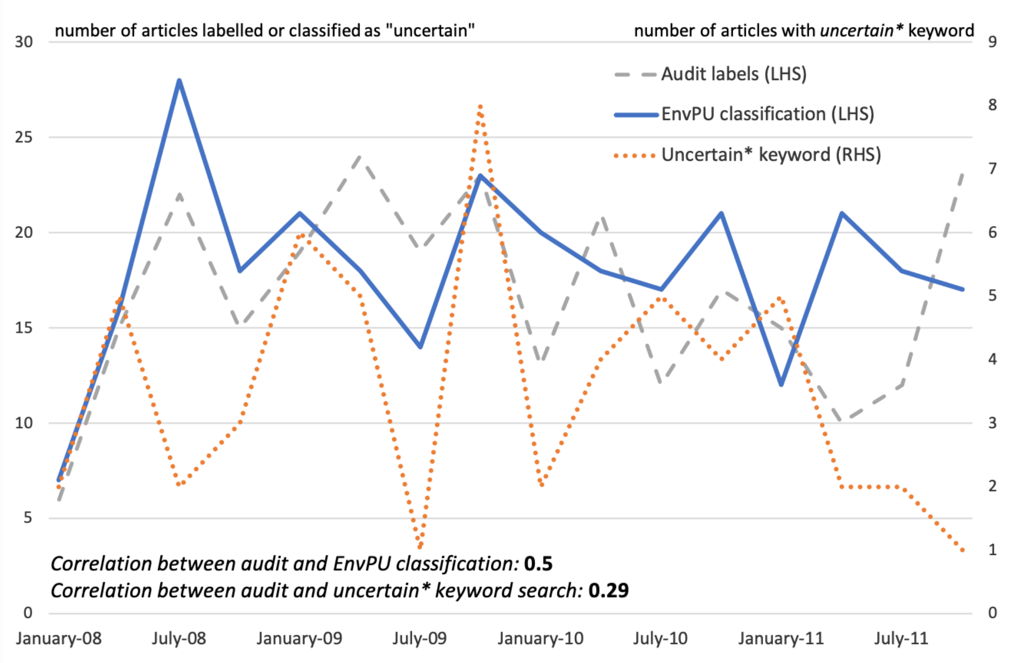
4.3.2 EnvPU human audit

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

By providing us with a human-based index that we can compare to the machine-based indices (i.e., SVM and *uncertain** approach), we aim to verify that our computer-based index does not miss any significant change in the level of policy uncertainty. To that end, we hired six human auditors to read and label a sub-sample of 925 articles randomly drawn from all the 14,158 published articles over the January 2008-December 2011 period (i.e., around 6.5 percent). We choose this period because it is a central period in terms of US environmental policy uncertainty, with Obama taking over as president and pushing a green policy agenda. Unaware of the classifier's label, the annotators manually labeled each article in pairs of two based on our codebook and continuous training we provided. Using the audit labels, we are able to verify the validity of our index and how it compares to the performance of the

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

Figure 4.5: The audit versus the EnvPU and uncertain* approach



*uncertain** approach.

Our human audit study reveals that human annotators agree with the SVM label 72 percent of the time. There is still a significant level of disagreement. However, this seems to be a feature of the task rather than a problem specific to machine learning. Indeed, when comparing the ten articles that are both in the SVM's training set and the auditors' set, we see that the two different groups of annotators for the training set and the human audit only agreed 60 percent of the time. Given the fact that these two groups were given the same codebook, it is fair to assume that some level of disagreement is unavoidable. Our algorithm, as well as any imaginable approach, will thus make mistakes. As a result, we focus more on testing whether humans and our algorithm pick up the same trends rather than on article-level accuracy.

We now investigate whether our SVM index picks up the same trends as our human-based index and how this compares to the *uncertain** keyword approach. Figure 4.5 plots these three EnvPU indices based on the number of articles identified as relevant on a quarterly basis. While the human- and SVM EnvPU indices capture the same overall trend of increasing uncertainty during 2008 followed by a small decrease, with an uptick in 2011, this trend is less visible when using the *uncertain** keyword approach. Furthermore, the correlation between the human-based and SVM index (0.5) is significantly higher than that between the keyword-based and the human-based index (0.29).

Finally, we test whether our machine learning algorithm makes systematic errors that are influenced by outside events such as the business cycle. To test this, we compute the difference between the number of articles labelled as uncertain by the human auditors and the algorithm and look at the correlation between this difference and a measure of economic growth (i.e., the quarter-on-quarter GDP growth from the U.S. Bureau of Economic Analysis). With a correlation coefficient of -0.13 over 2008-2011, we are confident that the errors of our classifier are not correlated with economic activity.[12]

In conclusion, using our audit as a benchmark we find that our machine-learning approach is able to capture similar trends in environmental policy uncertainty than humans and both approaches perform better than keyword-based approaches.

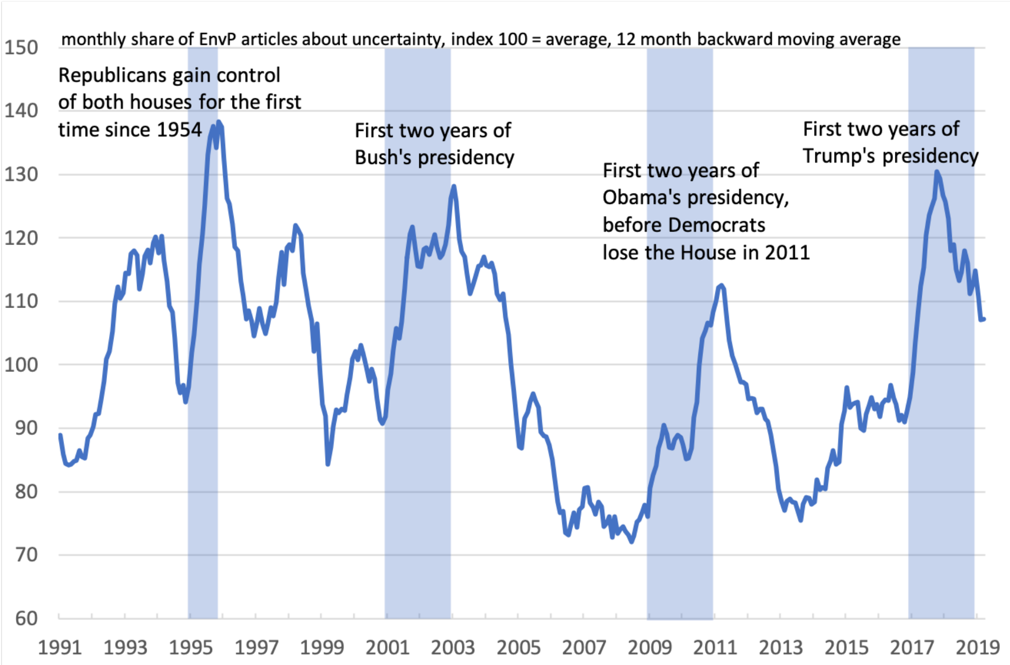
4.3.3 EnvPU and political cycles

The first striking feature of our index is that it seems to pick up during periods of transitions in US politics. Specifically, it spikes when Republicans retake control over important policy-making institutions. Figure 4.6 highlights these periods and plots a smoothed version of our index to showcase this stylized fact. The fact that policy uncertainty rises during transfers of power to Republicans is not particularly surprising. Indeed, the Republican policy agenda usually includes sweeping roll-backs of environmental regulations, causing a period of great uncertainty about the future state of environmental policy. We can see that the first spike in our index occurs between January and March 1995 when the Republicans take back both houses of Congress for the first time in 40 years and launch an attack on environmental regulations. Other notable high levels of policy uncertainty come about in 2003 when the Republicans control once again both houses and in 2017 when President Trump takes office with a clear agenda of rolling back Obama-era regulations.

Another interesting observation is that, after the initial period of high uncertainty when Republicans take power, our EnvPU index usually subsides at a low level. A notable illustration of this phenomenon can be found during the Trump era. Based on Figure 4.2, we can see that uncertainty stands at around 80, 20 percent below its average level, just before President Trump is elected and takes office at the end of 2016. EnvPU then nearly doubles to 160 in 2017, one of its highest level of the past three decades. However, once the dust has settled, policy uncertainty starts to creep back down, nearly reaching 80 once again in early 2019. This seems to suggest that a low but predictable level of environmental regulations, i.e., when Republicans are firmly in power, translates into a low level of environmental policy uncertainty. This is encouraging because it provides evidence that our index captures uncertainty and not simply the level/stringency of environmental policies.

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

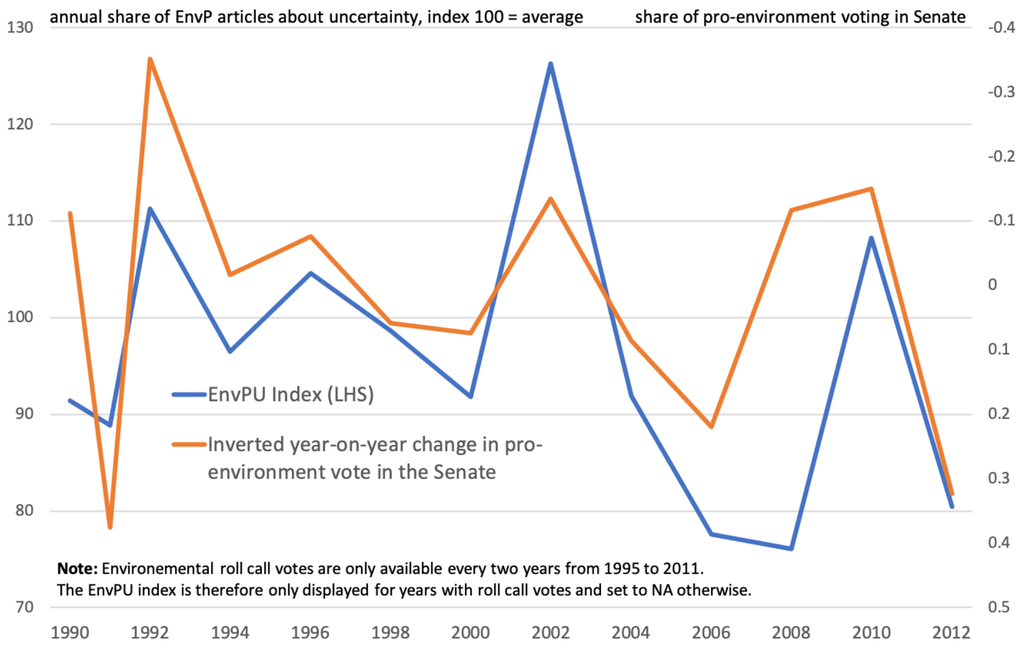
Figure 4.6: EnvPU during transitional periods in American politics, 12-month moving average



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

Figure 4.7 plots our EnvPU index at an annual frequency and the year-on-year change in the pro-environment votes in the Senate. We see that when the share of pro-environment votes drops in 1992-1993, 2002 or 2009-2010 the level of environmental policy uncertainty spikes. Increases in pro-environment votes also correlate with lower EnvPU. Overall, with a correlation coefficient of 0.6 over the years 1990 - 2013, our EnvPU index is associated with changes in the political composition of the American Congress. In particular, transfers of power to Republicans lead to spikes in environmental policy uncertainty.

Figure 4.7: EnvPU and environmental roll call votes



Descriptive evidence seems to indicate that our EnvPU is sensitive to political cycles. To investigate this relationship more formally, we now study the evolution of environmental policy uncertainty in the months before and after a presidential election.

Over our sample period of 1981-2019, we observe nine election cycles, starting with the 1984 presidential election between Reagan and Mondale up to the 2016 presidential election between Trump and Clinton.[13] We define each of the election cycles as the 22 months before and the 25 months after the election, as well as the month of the election itself. In a similar fashion to Baker et al. (2020), we then characterize the evolution of EnvPU in the months around the nine presidential elections by running the following regression:

$$ln(EnvPU_t) = \gamma_m + \gamma_c + \sum_{n=-6}^{6} \beta_n 1(ElectionMonth_{t-n} = 1) + \epsilon_t \tag{4.1}$$

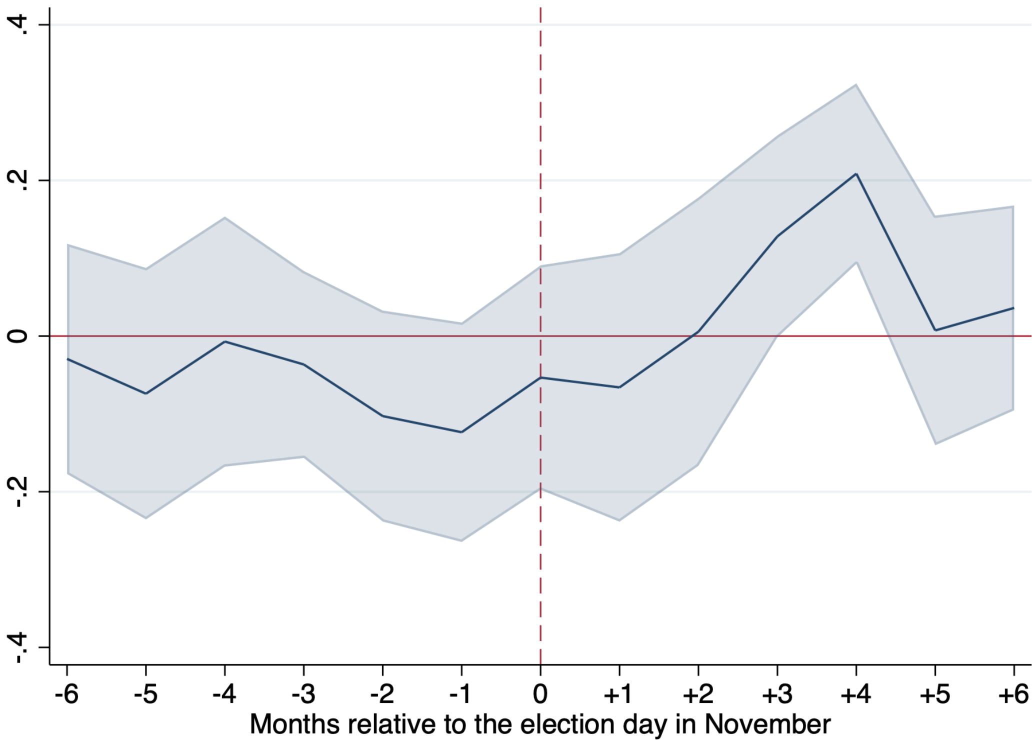
where t indexes the monthly dates, γ_m is a month fixed effect that deals with the potential seasonality in EnvPU and γ_c is an election cycle fixed effect. Each of the thirteen β_n coefficients captures the level of $ln(EnvPU)$ during the n months surrounding the election relative to the average level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

The estimates are in Column (1) of Table D.4 in Appendix D. To better visualize the behavior of EnvPU, Figure 4.8 displays the β_n coefficients for the six months before and after a typical

[13]Table D.3 in Appendix D details our nine election cycles.

presidential election. This Figure shows that policy uncertainty is slightly lower in the months directly before the election while it is around 21 percent higher four months after the election. The low level of environmental policy uncertainty or disagreement before presidential elections has already been documented by McAlexander and Urpelainen (2020). They show that both Republican and Democrat legislators are more likely to cast a pro-environment vote in the 60 days prior to an election. The rationale is that election time is when citizens pay the most attention to legislators, giving the latter incentives to engage in visibly pro-environment behavior.[14] Furthermore, the increase in EnvPU four months after the election is consistent with our analysis from the previous section. It is indeed typically during the first few months of their presidency that presidents announce which regulations they will rollback (usually the Republicans) and which regulations they will aim to implement (usually the Democrats). These announcements and the subsequent actions cause policy uncertainty to rise.

Figure 4.8: EnvPU and presidential elections



This figure presents the coefficients on dummies for six months before and after a presidential election using Equation (4.1) (i.e., Column (1) in Table D.4). The value of these coefficients reflects the level of EnvPU during each of these months relative to the rest of the sample. The shaded area depicts the 90% confidence interval.

To further explore the relationship between elections and EnvPU, we study whether elevated levels of polarization in environmental public opinion affect the behaviour of EnvPU around

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

elections. We determine whether elections in our sample are polarized using data from Kim and Urpelainen (2017) on the state-level differences in pro-environmental attitudes between Republicans and Democrats.[15] Both Figure D.2 and Table D.4 show that during polarized elections, the swings in EnvPU are more pronounced. Four months after a polarized election, environmental policy uncertainty is around 35 percent higher than during the rest of the election cycle. By contrast, none of the coefficients on the month dummies are significant when elections are not polarized.

Our EnvPU index is therefore sensitive to political cycles, especially when attitudes about the environment are polarized. As politics matter for our index, we now study how newspapers' partisan coverage might affect the EnvPU.

4.3.4 Political slant in newspapers

A newspaper-based measure of policy uncertainty is likely to be influenced by the political slant of the newspapers in our sample. Indeed, uncertainty is a rather subjective concept and conservative and liberal-leaning newspapers might overemphasize uncertainty when it is caused by their political opponents' action. To study whether political slant influences our index, we divide the newspapers in our sample into two groups based on whether they are more left or right leaning.[16]

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

We plot the EnvPU indices produced by the Liberal-leaning and Conservative-leaning newspapers in our index in Figure 4.9. An interesting observation is that, while the indices are very similar most of the time, they tend to diverge during periods of high political uncertainty. Left-leaning journalists describes environmental policy in very uncertain terms when Republicans suddenly regain control of the political apparatus while the more right-leaning newspapers are less affected. The most striking example of this occurs in the last years of our index. On the one hand, right-leaning newspapers report the same level of uncertainty when Trump takes office than during the last two years of Obama's presidency. If anything, they describe environmental policy as being less uncertain under Trump. Indeed, at the start of 2019, the

[15] We aggregate this value at the federal level for each race year in the House of Representatives. An election cycle is polarized if our measure of polarization during the year of the presidential election is above its average value over 1980-2012. The presidential election in 2016 is assumed to have been polarized. Table D.3 indicates which election cycles are classified as polarized.

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

level of uncertainty reported by conservative newspapers is at one of its lowest level of the past three decades. On the other hand, left-leaning newspapers report a near doubling of policy uncertainty when Trump is elected, to its second highest level of the past three decades. Moreover, policy uncertainty remains high throughout Trump's early presidency. A similar scenario can be observed in 1995 when Republicans regain full control of Congress.

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

Overall, the fact that political slant skews the reporting of policy uncertainty is interesting but does not undermine our index. Indeed, our sample of newspapers is well balanced between liberal and conservative prints.[17] As a result, our EnvPU index evens out the biases from each side.

Figure 4.9 also highlights the need to have a geographically diverse set of newspapers. Indeed, in the 2011-2012 period, the Dallas Morning News and Houston Chronicles report a high level of policy uncertainty while other newspapers are less affected. This is due to the fact that the source of this uncertainty, a fight between Texas and the EPA over clean air rules, happens in the home state of these newspapers. Journalists that are directly impacted or exposed to a policy dispute might overemphasize the sense of uncertainty.

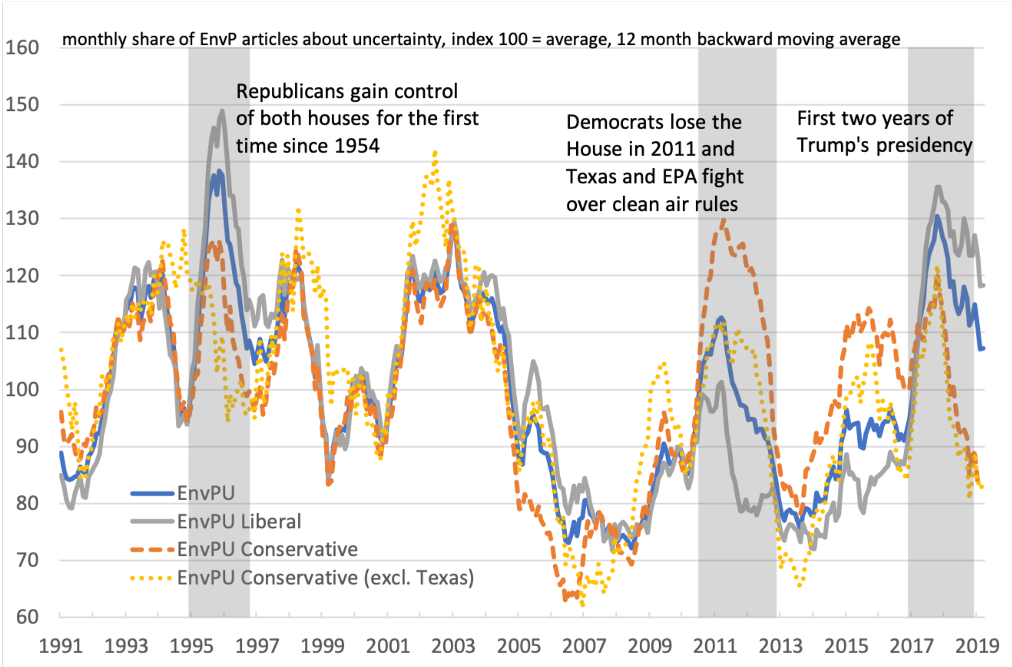
4.4 Environmental policy uncertainty and low-carbon investments

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

Conceptually, we expect that a rise in environmental policy uncertainty is always bad news for low-carbon investments, as these tend to be heavily reliant on public policies. Independent of the current level of policy stringency, uncertainty about the future regulatory framework may threaten market opportunities and therefore the profitability of clean investments. We thus expect environmental policy uncertainty to be negatively associated with venture capital funding for startups engaged in the low-carbon economy. In financial markets, we expect our EnvPU index to be associated with higher volatility for stock returns of firms active in low-carbon activities.

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

Figure 4.9: EnvPU according to liberal and conservative media



Causality is challenging, as many omitted variables could affect both environmental policy uncertainty and low-carbon investments, in particular other forms of uncertainty related for instance to technological or climate uncertainty. Yet, two aspect mitigate these concerns. First, we expect that the correlation between these omitted types of uncertainty and our EnvPU is small – motivated by the fact that we explicitly labelled non-policy-related uncertainty as not relevant in our labeling exercise. Then, part of the policy uncertainty captured by our index is driven by exogenous political cycles and the timing of elections, which are unaffected by low-carbon markets.

Our identification strategy relies on a differentiation of firms according to their exposure to environmental policy uncertainty. Specifically, we expect to find a stronger empirical association between our EnvPU index and investments in firms active in the low-carbon economy – those more exposed to environmental policy uncertainty – compared to other firms, which should remain unaffected. We also control for many confounding factors via fixed effects and additional variables. In particular, we control for the salience and level of environmental policy, which allows us to identify the effect of policy uncertainty for a given level of perceived policy stringency. As an additional step, we complement firm-level estimations with VAR models to illustrate the dynamic relationship between our EnvPU index and low-carbon investments at the aggregate level.

4.4.1 Firm-level estimations

VC investments across industries

We first investigate the impact of environmental policy uncertainty, measured by our EnvPU index, on the probability that a startup will receive VC funding. If our index is valid, VC funding should be more responsive to changes in our EnvPU index for firms and startups more exposed to environmental policy uncertainty. To test this, we create three industry-categories based on their exposure to EnvPU. First, we have non-cleantech startups. Uncertainty about the state of future environmental policy should have little direct impact on the fate of startups in sectors like ICT or medtech. Second, we have non-energy cleantech startups (e.g., pollution filters or recycling). As the business model of these startups depend, in part, on environmental policy support, policy uncertainty has a direct impact on their investment plans. Third, and most exposed to EnvPU, we have clean energy startups (renewable energy such as wind and solar). The literature has shown that firms with a higher degree of investment irreversibility tend to be more sensitive to policy uncertainty (Guien and Ion, 2016). The inability to predict future policy incentives makes it particularly risky to invest in capital-intensive assets that will pay off over long time horizons (e.g., factories, complex machinery). Clean energy startups require more capital and time to commercialize than other clean industries because of the need to manufacture and deploy new solar and wind technology (Nanda et al., 2015; Gaddy et al., 2017; Popp, 2017). In line with the literature, we expect an increase in the EnvPU index to be negatively associated with VC financing for cleantech startups (Tian and Ye, 2017), especially for clean energy startups because of their higher exposure to EnvPU. By contrast, we do not expect much of a relationship between our EnvPU index and VC financing for non-cleantech startups.[18]

To test these predictions, we obtain data on VC funding rounds between January 1998 and March 2019 for US startups from the Crunchbase database and aggregate these funding rounds into a firm-quarter panel dataset.[19] We also extract the firm's industry and founding date as well as all the information related to the funding rounds (i.e. date, amount, series) from Crunchbase. Firms are only included in our analysis when they are active. This means that we remove any firm-quarter observations that occur before a startup's founding date. We also remove observations of startups no longer seeking funding, either because they have gone bankrupt or because they are mature enough not to require VC funding.[20] In order to control for the level of media attention to environmental policy, we use our Environmental Policy Index (EnvP) described in Section 4.2.1. Finally, we include GDP data from the U.S. Bureau

[18] It is possible that our index marginally captures general sentiments of uncertainty that affect all startups, but even then the effect of EnvPU should be concentrated on cleantech startups.

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

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

of Economic Analysis, the Federal Reserve Fed Funds rate and the West Texas Intermediate crude oil spot price. Excluding firm-quarter observations containing any missing information — including those where the firm is classified as inactive — we obtain 1,056,221 firm-quarter observations on 35,704 unique startup firms. Table D.5 in Appendix D provides summary statistics of the variables in our sample.

Cleantech startups belong to Crunchbase's 'Sustainability' industry group and represent 4 percent of overall VC deals, while clean-energy startups in clean energy, battery, renewable energy, wind energy, energy storage and solar industries represent only 2.4 percent of all VC deals.[21] We estimate whether startups that are classified as cleantech or clean energy are significantly more responsive to our EnvPU index than startups in other sectors using ordinary least squares (OLS) as follows:

$$\begin{aligned} VC_{i,t+s} = &\alpha + \beta_1 EnvPU_t + \beta_2 EnvPU_t \cdot Cleantech_i + \beta_3 Controls_{i,t} \\ &+ \beta_4 TimeTrend_t + \gamma_{quarter/year/industry/state/series} + \epsilon_{i,t} \end{aligned} \tag{4.2}$$

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

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

We also include a set of variables and fixed effects to absorb variation that is unrelated to environmental policy uncertainty but may nonetheless affect our results, including, firm i's age as well as a time trend, and in some specifications an industry time trend. We also use firm, quarter, year and series funding round fixed effects.[22] The quarter fixed effects

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

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

are used to account for seasonality in the data.[23] The firm fixed effect control for firm-level unobservables such as quality. The other fixed effects also allow us to control for unobserved variables common to all startups in a given year or funding round. Finally, we cluster standard errors at the startup firm level to correct for potential serial correlations in the error term.

Table 4.3 presents the results of our regressions using Equation 4.2. Columns (1) and (2) use the probability of getting funded in the next quarter (Q + 1) as the dependent variable. In column (1), we only differentiate between cleantech and non-cleantech startups. We can see that a rise in our EnvPU index is associated with a reduced probability of receiving funding in the next quarter for cleantech startups. The positive coefficient for non-cleantech startups could indicate that a rise in EnvPU makes non-cleantech sectors more attractive to VCs. However, this coefficient is relatively small in size, around one forth the size of the cleantech interaction term, which underlines that the effect is concentrated on cleantech startups. In column (2), we separate between clean energy and other cleantech startups. Doing so, we find that policy uncertainty has a stronger adverse impact on clean energy startups than on the rest of the cleantech startups.[24] To illustrate the size of the effect on clean energy investments, a one-standard deviation (sd) increase in environmental policy uncertainty from one quarter to the next would decrease the probability of receiving funding by 0.24 percentage points. While this might seem small, the average probability that a clean energy startup will be funded next quarter in our sample is only 6.4 percent. Therefore a one-sd increase in environmental policy uncertainty is actually associated with a 4 percent decrease in a clean energy startup's probability of receiving funding next quarter.

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

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

To further test the validity of our EnvPU index, we do the same analysis with alternative policy uncertainty indices. First, we estimate the relationship between investments and the naive environmental policy uncertainty index based on the uncertain* keyword search.

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

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

Table 4.3: Baseline results: Relationship between environmental policy uncertainty and VC investment in cleantech

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

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

Columns (1) and (2) of Table 4.4 shows that the naive index has a weaker relationship with cleantech than the EnvPU index. Indeed, in Column (1) both cleantech and non-cleantech startups' probability of securing funding are negatively associated with the naive index. More importantly, there are no specific cleantech effect when looking at the amount of funding in column (2). In column (3) we test whether Baker et al. (2016)'s Economic Policy Uncertainty (EPU) index has a particular effect on cleantech startups. We do so because the EPU should have an effect on all startups, irrespective of their industry, but no specific cleantech effect. This is what we find, a one standard deviation increase in the EPU is associated with a 0.27 percentage point decrease in the probability of getting funded, with no cleantech effect. The fact that the EnvPU index has a specific effect on cleantech but the EPU does not reinforces our confidence in the EnvPU index.

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

Firm-level stock volatility

Next, we consider the relationship between environmental policy uncertainty as measured by our EnvPU index and firm-level stock volatility. We measure exposure to environmental policy based on the average green revenue share at the firm level, as we expect firms with especially high green revenue shares to be more dependent on environmental policies in terms of business outlook. Two related aspects are important here. First, green firms often compete with highly profitable dirty firms in the same market (think of energy), which offer an attractive outside option for investors in the face of environmental policy uncertainty. Second, as a result, since policy support is often what facilitates investment in these green firms in the first place, uncertainty around such policies is likely to lead to higher stock volatility as investors find it harder to assess the future profitability of green firms.

In line with previous literature which has shown that policy uncertainty harms investments and leads to higher stock market volatility of exposed firms (Baker et al., 2016; Pástor and Veronesi, 2012), we expect investors to pull back from green investments when their future returns become uncertain. Pástor and Veronesi (2012) predict that stock volatility rises when policy changes are announced, because the new policy's impact uncertainty increases the volatility of agents' stochastic discount factor. Uncertainty in the stock market usually translates into higher discount rates because agents care less about the future if they cannot count on future

Table 4.4: Robustness: Relationship between alternative policy uncertainty indices and VC investment in cleantech

	(1) Funded (Q+1)	(2) Amount (Q+1)	(3) Funded (Q+1)
EnvP index	-0.00272***	-0.0159	-0.00312***
	(0.000730)	(0.0176)	(0.000728)
EnvP x Cleantech excl. Energy	0.00271***	0.0428	0.00280***
	(0.000973)	(0.0321)	(0.00106)
EnvP x Clean Energy	0.00480***	0.0642***	0.00483***
	(0.00105)	(0.0249)	(0.00112)
Naive index	-0.00189***	-0.00931	
	(0.000360)	(0.00799)	
Naive x Cleantech excl. Energy	-0.00190*	0.0325	
	(0.00101)	(0.0250)	
Naive x Clean Energy	-0.00215**	0.00928	
	(0.000978)	(0.0227)	
EPU index			-0.00281***
			(0.000678)
EPU x Cleantech excl. Energy			-0.000556
			(0.000868)
EPU x Clean Energy			-0.000462
			(0.000935)
Firm FE	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Industry-time trend	Yes	Yes	Yes
Series FE	Yes	Yes	Yes
Observations	1056221	57319	1056221
Firms	35637	28297	35637
R^2	0.006	0.119	0.006

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

profits. Volatility here could mean that agents have different beliefs about the future. In Pástor and Veronesi (2013), stock returns are more volatile when policy uncertainty is higher.

We collect the daily stock price indexes for a sample of 438 firms across 9 industries[25] listed on the US stock exchange from January 2008 to March 2019 from Datastream. We also compute daily continuously compounded log returns at the firm level as $r_{i,t} = \ln\left(\frac{p_{i,t}}{p_{i,t-1}}\right)$ and then monthly annualized volatility at the firm level as $\sqrt{252} \times \sigma_{r_{i,t}}$.

Since our EnvP and EnvPU indices may be endogenously affected by market activity or anticipated, we follow Brogaard and Detzel (2015) and consider 'innovations' in both indices by extracting the residuals from an AR(7) and AR(3) model of our monthly series of EnvP and EnvPU, respectively. Standard tests confirm that both series are white noise and have no autocorrelation.[26] We standardize these measures to have unit standard deviation.

We capture firm-level exposure to media attention on environmental policy by green revenue shares from FTSE Russell.[27] In some years, the green revenue share is estimated and a range instead of a point estimate is provided. In this case, we always choose the lower bound of the interval as our estimate of firm green revenue shares. The average firm in our sample has a green revenue share of about 23 percent. 'Energy', 'Utilities' and 'Consumer Stapes' industries have the highest green revenue share.

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

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

where $\log(\sigma_r)$ is the natural logarithm of the annualized realized volatility of the continuously compounded log returns at the firm level, Green Revenue share$_{i,t=y}$ is either the AVG annual green revenue share at the firm level or the pre-sample green revenue share at the firm level. We test both to make sure that our EnvPU exposure measure is not endogenous to EnvPU itself as firms may choose to increase or decrease their green revenue share depending on the environmental policy outlook as well as choose to enter the sample at a later stage due to strategic considerations. The γ's are firm and month fixed effects, respectively. We provide summary statistics of all variables used in the regression in Table D.6 in Appendix D.

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

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

[27] Green revenues are captured by 10 sectors, 64 sub-sectors and 133 micro sectors. The 10 sectors are: energy generation, energy management and efficiency, energy equipment, environmental resources, environmental support services, food & agriculture, transport equipment, transport solutions, water infrastructure & technology, and waste & pollution control. For further details, refer to the FTSE Green Revenue Classification System.

Table 4.5 shows our results. Environmental policy uncertainty has a significant and sizable positive effect on stock volatility of the greenest firms, regardless of whether we use pre-sample GR shares or average GR shares. Quantitatively, a one-standard deviation EnvPU innovation leads to a differential increase in volatility for a firm one standard deviation above the mean in terms of average green revenue share of about 0.5 percent since $\exp(0.005) \approx 1.005$. This effect is significant at the 5 percent level and robust to including industry fixed effects, indicating that there is meaningful within-industry variation in the green revenue share (we do not show this here because the coefficients stay the same). Moreover, in column (2) we compare the top 10 percent firms in terms of their average green revenue share with the bottom 10 percent and find that their stock volatility rises by 1.3 percent in response to an EnvPU shock. In addition, when using the pre-sample green revenue share as an exposure measure as in column (3), a one-standard deviation EnvPU innovation leads to a volatility increase of about 1 percent. This effect is significant at the 1 percent level and also robust to including industry fixed effects as well as an industry-month time trend. Moreover, columns (4)-(5) show that the effect of the EnvPU remains even when controlling for general Economic Policy Uncertainty (EPU). Finally, in columns (6)-(7) we find no effect when using the naive EnvPU index, further underlining the value-added of our text mining based EnvPU index.

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

Table 4.5: Estimation results - EnvPU and Stock Volatility

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

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

4.4.2 Dynamic estimations at the aggregate level

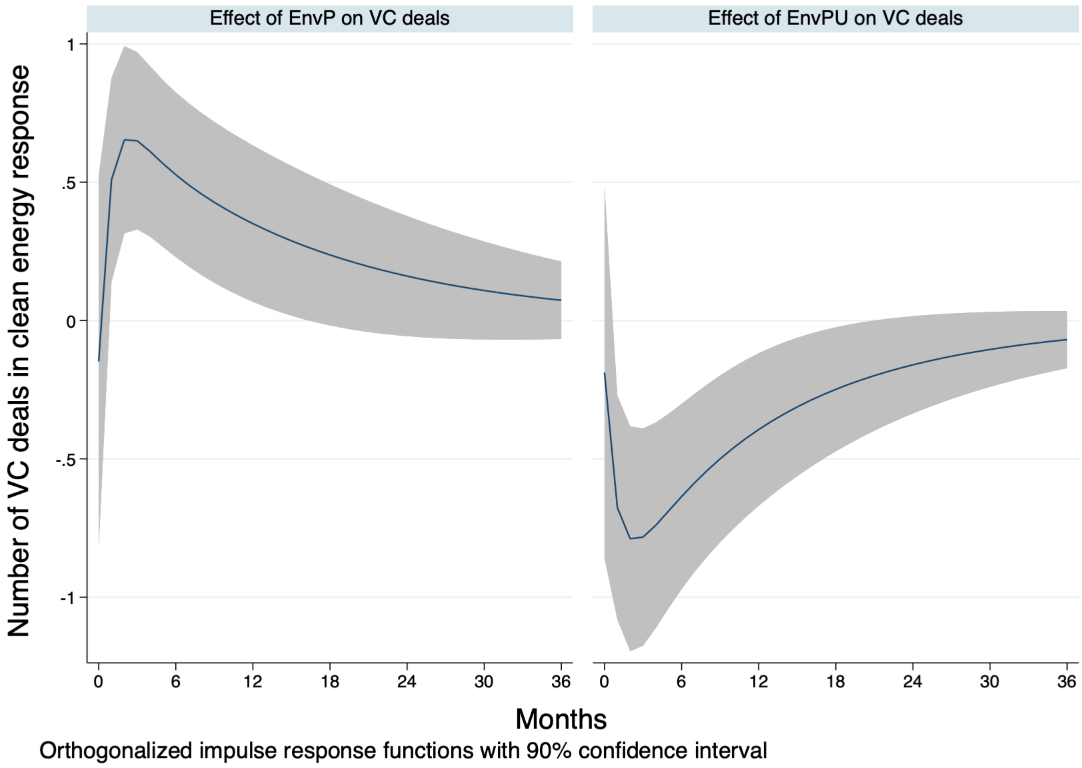
Aggregate cleantech venture capital deals

We now consider the dynamic relationship between policy uncertainty and aggregate venture capital (VC) activity in clean energy technologies. To study this relationship, we use i3 Cleantech Group's data on early-stage financing of cleantech startups, which offers consistent coverage over the past two decades. This database provides information on 11,620 early-stage cleantech deals (seed, series A, series B and growth equity) in the U.S. tracked over time by the Cleantech Group. We extract data on the monthly number of VC deals in the 'energy & power' classification (which includes clean energy generation, efficiency, storage and infrastructure) from January 1998 to March 2019. We focus on VC deals involving clean energy startups because, as discussed in Section 4.4.1, these are the investments that should be most exposed to uncertainty in environmental policy.

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

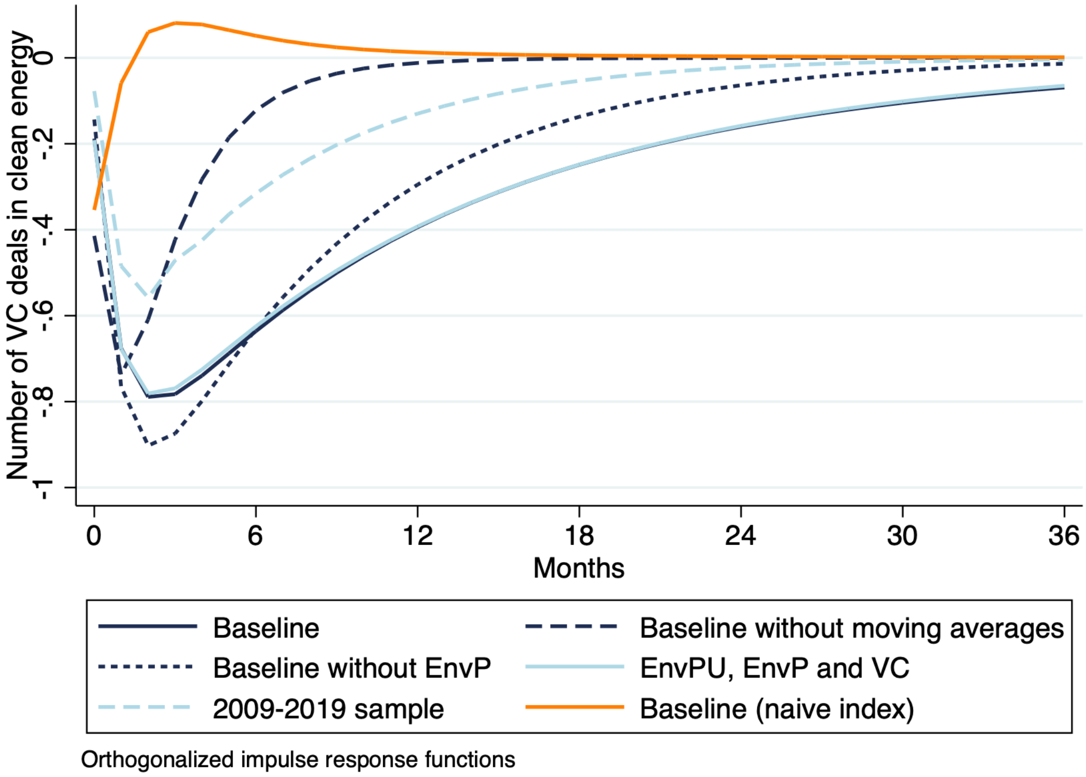
Figure 4.10 displays the orthogonalized impulse response functions of the number of VC deals in clean energy to both a shock to environmental policy uncertainty and a shock to environmental policy. In the right panel, we see that a one standard deviation increase in environmental policy uncertainty is associated with between 0.5 and 0.8 fewer VC deals in clean energy during the first year after the shock. Conversely, a one standard deviation increase in environmental policy is associated with around 0.5 more VC deals in clean energy (see left panel). While the effect of policy uncertainty is moderate in size, losing half a VC deal still represents a sizable 4.2 percent decrease in the average monthly number of VC deals in clean energy in our sample (i.e. 15.6 between January 1998 and March 2019).

Figure 4.10: Estimated effect of a one-sd shock in EnvP and EnvPU on the number of clean energy venture capital deals. Impulse response functions to our news-based EnvP and EnvPU indices. The policy indices are smoothed using a three-month backward-looking moving average.



We test the robustness of our results to varying specifications and to using the naive environmental policy uncertainty index instead of our EnvPU index in Figure 4.11. The negative relationship between EnvPU and VC investments in clean energy holds across specifications. Our baseline response is very similar to the response from a VAR model without our EnvP index, as well as the response without including our controls. Moreover, changing the sample to only cover the post-Great Financial Crisis (July 2009) period or not using moving averages to smooth our indices, does not fundamentally change our results. Interestingly, Figure 4.11 also shows that the naive index does not display any meaningful dynamic relationship with VC investments, which confirms our previous result that our EnvPU index is more meaningfully related to cleantech investments than one based on a more naive keyword approach.

Figure 4.11: Clean energy venture capital deals responses to EnvPU Shock (IRFs), under alternative specifications and samples.



Aggregate clean energy stocks

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

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

Figure 4.12: Estimated effect of a one-sd EnvPU shock on the volatility of the AuM of the PBW ETF. Impulse response functions of our news-based EnvPU index. The policy indices are smoothed using a three-month backward-looking moving average.

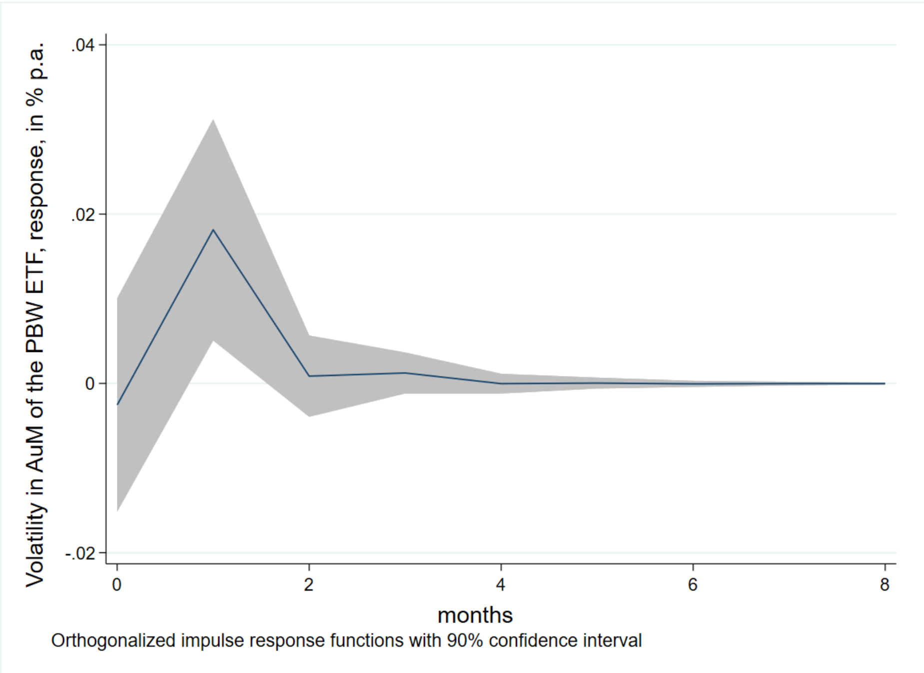
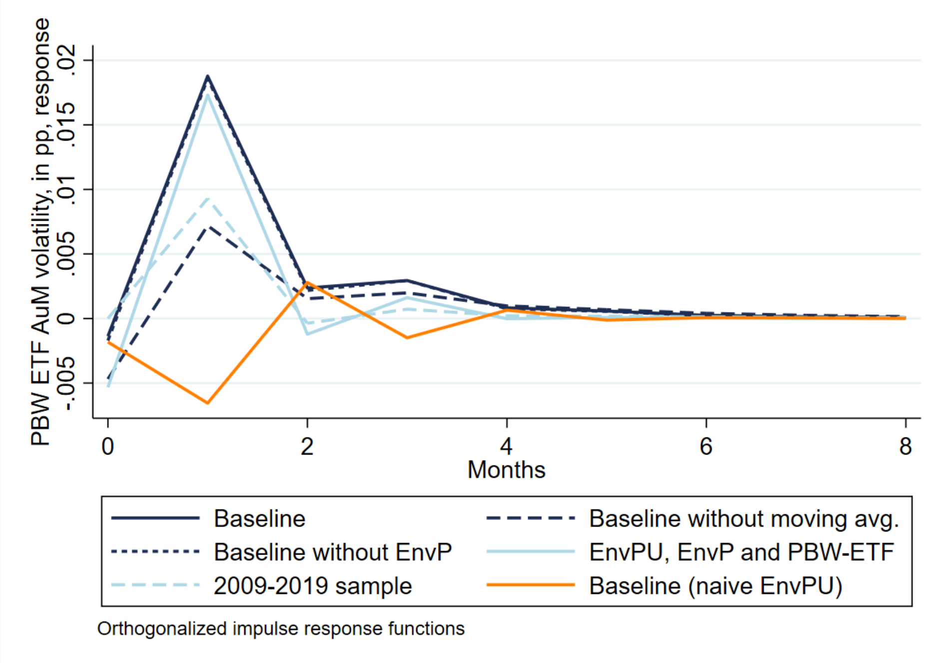


Figure 4.12 shows our result. A one-standard deviation shock to the growth rate of the EnvPU index leads to an increase in the volatility of the AuM of the PBW ETF of 0.02 percentage points per annum one month after the shock. One reason for this lag may be that ETF investors are largely retail investors who tend to be more passive and less sophisticated in their investment decisions than institutional investors. Therefore, the link between policy uncertainty and investment decision may have more of a medium-term nature.

Figure 4.13 contains our robustness checks. We test a specification without moving average transformation of our EnvPU and EnvP indices, one without the EnvP index, one with only our main variables of interest, one limiting the sample to the post GFC period (July 2009-March 2019) as well as one using the naive EnvPU index instead of our text mining index. Evidently, the results are weaker when using the non-smoothed policy indices. This is however not surprising given the rationale explained above. Moreover, the effect substantially weakens after the crisis. There is also no effect when using the naive EnvPU index, further underlining the usefulness of our EnvPU index.

Figure 4.13: Volatility responses of the AuM of the PBW ETF to EnvPU Shock (IRFs), under alternative specifications and samples.



4.5 Conclusion

A predictable regulatory framework is key to mobilize investments and financial flows towards the low-carbon economy. We apply text-mining techniques on ten leading US newspapers to construct a new index of environmental policy uncertainty – the EnvPU index – over the 1990-2019 period. Our index captures the monthly share of policy uncertainty news in environmental policy articles. We find that about one-third of environmental policy news report about policy uncertainty, suggesting that the inability to predict how future environmental regulations will unfold is a pervasive attribute of environmental policy discourses. Our EnvPU index correctly captures important spikes in policy uncertainty in the history of US environmental policy, such as the collapse of the national cap-and-trade policy proposals in 2010 and the environmental policy rollbacks under the Trump's administration as of 2017. We discuss how our novel methodology based on supervised machine learning algorithms outperforms other kewyword-based approaches and we conduct an additional accuracy check using a human audit study. In addition, we show that our index relates to transitions in political and electoral cycles in a meaningful manner. Finally, we address concerns about potential bias due to newspapers partisan coverage.

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

Policy changes and reforms are an inherent part of democratic policymaking. Election cycles contribute to important fluctuations in the environmental regulatory framework, due to recurrent precedence of short-term economic concerns over long-term environmental ones. In addition, environmental regulations are typically based on long-term considerations, with science and technological progress playing an important role, and policies need to be adapted as new information arrives. A challenge for effective policymaking is thus to allow for some necessary flexibility, while still providing a stable and predictable policy framework. Several options exist to alleviate the trade-off between flexibility and predictability. Gawel and Lehmann (2019) contrast for instance flexibility by policy design – where policy incentives such as renewable subsidies adjust automatically and in a predictable way to varying market conditions, such as electricity prices and technological costs – and flexibility by periodic policy adjustments – where policies are revised according to pre-announced cycles. More ambitious proposals for safeguarding long-term environmental and climate policy action from the volatile politics of election cycles recommend creating new institutions built on independent expertise and a long-term perspective. These new institutions would diminish rather than exacerbate policy uncertainty, thereby providing the right conditions for long-term investment and innovation (Aghion et al., 2013).

Conclusion

Fighting climate change is one of the greatest challenges of the 21st century, if not the greatest (Lopez-Claros et al., 2020). However, the scope of investments in clean technologies does not yet match this sense of urgency. This thesis offers findings that could help direct funding towards clean technologies and improve the efficiency of such funding.

In Chapter 1, we show that entrepreneurial programs like venture competitions can improve the allocation of funding to very early-stage startups. The certification of quality they provide accelerates the termination of low-quality startups and improves external funding opportunities for high-quality startups. However, for this certification effect to materialize, entrepreneurs and investors need to trust the certification to be informative. Fortunately, clean technologies are like other science-based sectors in the sense that judges' certification is likely to be meaningful. Chapter 2 provides insights into the recent VC boom in clean technologies (Bullard, 2021). Sectors like electric vehicles and plant-based meats are now attracting a lot of VC funding thanks to dynamic demand for these cleantech goods and the outsized returns of front-runners like Tesla or Impossible Foods. However, in the absence of carbon pricing, clean energy startups could continue to struggle to attract VC. To support clean energy startups, governments should implement demand-side policies that make investing in clean energy more viable rather than investing in startups bound to struggle through the valleys of death. With demand-side policies in place, governments could then turn to funding firms in cleantech markets where product differentiation is difficult, as such firms are least likely to attract private sector support.

The last two chapters offer both novel findings and data that improve the understanding of the effect of environmental policy on investments in clean technologies. Our EnvP index and its sub-components, developed in Chapter 3, can help policymakers and researchers measure the salience of various environmental policy topics in the United States. As our results show, the EnvP index displays a meaningful empirical association with investments in clean technologies. This finding suggests that news on environmental policy contains relevant information for investors. Policymakers could, therefore, leverage their capacity to shape investors' expectations and behavior by communicating audibly about their environmental policy agenda. Chapter 4 provides a valuable indicator of environmental policy uncertainty based on recent developments in text-as-data methods. This EnvPU index can help policymakers gauge how uncertain investors are about the state of future regulations. We also provide evidence that environmental policy uncertainty weighs on private investments. Policymakers should, therefore, ensure that their actions diminish rather than exacerbate policy uncertainty, for example, by implementing clear and predictable policy frameworks.

The questions explored in this thesis open the way for several future studies on the financing of young and innovative cleantech startups. In Chapter 1, we find, like Scott et al. (2020), that experts find it harder to accurately assess the potential of startups in ICT than science-based startups. It would be interesting to explore further the reasons behind this phenomenon (i.e., is there a lack of expertise in ICT on competitions' juries, or are ICT startups inherently harder to evaluate?). Chapter 2 shows evidence that clean energy startups have a lower potential for outsized returns than sectors like ICT and biotech. While we offer some explanations, it would be interesting to explore further why we do not observe outsized winners in clean energy and whether mechanisms other than demand-side policies can be implemented to make investing in young clean energy firms more appealing to VCs. Finally, Chapters 3 and 4 offer novel indicators of environmental policy and environmental policy uncertainty that can be used in future research. Moreover, they further validate the use of text-as-data methods in fields like environmental economics, which will hopefully encourage other researchers to use these valuable new tools.

Bibliography

Acemoglu, D. and Cao, D. V. (2010). Innovation by Entrants and Incumbents. *NBER WP 16411*.

Aghion, P., Besley, T., Browne, J., Caselli, F., Lambert, R., Lomax, R., Pissarides, C., Stern, N., and Van Reenen, J. (2013). Investing for prosperity: Skills, infrastructure and innovation. Technical report.

Albrizio, S., Kozluk, T., and Zipperer, V. (2017). Environmental policies and productivity growth: Evidence across industries and firms. *Journal of Environmental Economics and Management*, 81:209–226.

Amit, R., Glosten, L., and Muller, E. (1990). Entrepreneurial ability, venture investments, and risk sharing. *Management science*, 36(10):1233–1246.

Angrist, J. D. (2001). Estimation of Limited Dependent Variable Models With Dummy Endogenous Regressors. *Journal of Business & Economic Statistics*, 19(1):2–28.

Arrow, K. J. (1971). The economic implications of learning by doing. In *Readings in the Theory of Growth*, pages 131–149. Springer.

Åstebro, T. and Gerchak, Y. (2001). Profitable advice: The value of information provided by Canada's inventor's assistance program. *Economics of Innovation and New Technology*, 10(1):45–72.

Audretsch, D. B., Keilbach, M. C., and Lehmann, E. E. (2006). *Entrepreneurship and economic growth*. Oxford University Press.

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

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

Bakkensen, L. and Barrage, L. (2021). Climate shocks, cyclones and economic growth: Bridging the micro-macro gap. NBER Working Paper No 24893.

Barnett, M. (2020). The estimated impacts of climate policy risk.

Baron, D. P. (2006). Persistent media bias. *Journal of Public Economics*, 90(1-2):1–36.

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

Baylis, P. (2020). Temperature and temperament: Evidence from Twitter. *Journal of Public Economics*, 184:104–161.

Bernanke, B. (1983). Irreversibility, uncertainty and cyclical investment. *The Quarterly Journal of Economics*, 98(1):85–106.

Bernstein, S., Korteweg, A., and Laws, K. (2017). Attracting Early-Stage Investors: Evidence from a Randomized Field Experiment. *Journal of Finance*, 72(2):509–538.

Blei, D., Ng, A., and Jordan, M. (2003). Latent Dirichlet Allocation: Extracting Topics from Software Engineering Data. *Journal of Machine Learning Research*, 3:993–1022.

Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica*, 77(3):623–685.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., and Terry, S. J. (2018). Really Uncertain Business Cycles. *Econometrica*, 86(3):1031–1065.

Blyth, W., Bradley, R., Bunn, D., Clarke, C., Wilson, T., and Yang, M. (2007). Investment Risks under Uncertain Climate Change Policy. *Energy Policy*, 35(11):5766–5773.

Bohr, J. (2020). Reporting on climate change: A computational analysis of U.S. newspapers and sources of bias, 1997–2017. *Global Environmental Change*, 61:102–038.

Bonaime, A., Gulen, H., and Ion, M. (2018). Does policy uncertainty affect mergers and acquisitions? *Journal of Financial Economics*, 129(3):531–558.

Botta, E. and Kozluk, T. (2014). Measuring environmental policy stringency in OECD countries: A composite index approach. *OECD Economic Department Working Papers*, (1177).

Boykoff, M. T. and Boykoff, J. M. (2007). Climate change and journalistic norms: A case-study of US mass-media coverage. *Geoforum*, 38(6):1190–1204.

Bretschger, L. and Soretz, S. (2018). Stranded Assets: How Policy Uncertainty Affects Capital, Growth, and the Environment.

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

Brunel, C. and Levinson, A. (2016). Measuring the stringency of environmental regulations. *Review of Environmental Economics and Policy*, 10(1):47–67.

Brüggemann, M. and Engesser, S. (2017). Beyond false balance: How interpretive journalism shapes media coverage of climate change. *Global Environmental Change*, 42:58–67.

Bullard, N. (2021). Venture Capital Firms Are Fighting to Throw Money at Cleantech.

Bumpus, A. and Comello, S. (2017). Emerging clean energy technology investment trends. *Nature Climate Change*, 7(6):382–385.

Bybee, L., Kelly, B. T., Manela, A., and Xiu, D. (2020). The structure of economic news.

Cai, Y. and Lontzek, T. S. (2018). The Social Cost of Carbon with Economic and Climate Risks. *Journal of Political Economy*, 127(6):2684–2734. Publisher: The University of Chicago Press.

Carpenter, R. E. and Petersen, B. C. (2002). Is the growth of small firms constrained by internal finance? *The Review of Economics and Statistics*, 84(2):298–309.

Clementi, G. L. and Hopenhayn, H. A. (2006). A theory of financing constraints and firm dynamics. *Quarterly Journal of Economics*, 121(1):229–265.

Clingingsmith, D. and Shane, S. (2017). Let Others Go First : How Pitch Order Affects Investor Interest in Elevator Pitches. (September):1–33.

Cole, M. A. and Fredriksson, P. G. (2009). Institutionalized pollution havens. *Ecological Economics*, 68(4):1239–1256.

Congress (2017). Risky Business: The DOE Loan Guarantee Program.

Connelly, B. L., Tihanyi, L., Crook, T. R., and Gangloff, K. A. (2014). Tournament Theory: Thirty Years of Contests and Competitions. *Journal of Management*, 40(1):16–47.

Consoli, S., Barbaglia, L., and Manzan, S. (2021). Fine-grained, aspect-based sentiment analysis on economic and financial lexicon.

Dahal, B., Kumar, S. A. P., and Li, Z. (2019). Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining*, 9(1):24.

Dalle, J.-M., Den Besten, M., and Menon, C. (2017). Using Crunchbase for economic and managerial research Using Crunchbase for Economic and Managerial Research. *OECD Science, Technology and Industry Working Papers*, No. 2017/2.

Dasgupta, S., Laplante, B., and Mamingi, N. (2001). Pollution and capital markets in developing countries. *Journal of Environmental Economics and management*, 42(3):310–335.

Davenport, C. and Samuelsohn, D. (2010). Dems pull plug on climate bill.

Denny, M. J. and Spirling, A. (2018). Text preprocessing for unsupervised learning: Why it matters, when it misleads, and what to do about it. *Political Analysis*, 26(2):168–189.

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

Dorsey, J. (2019). Waiting on the Courts: Effects of Policy Uncertainty on Pollution and Investment. *SSRN Electronic Journal*.

Dugoua, E., Dumas, M., and Noailly, J. (2021). Text-as-data in environmental economics and policy. *Review of Environmental Economics and Policy*.

Durnev, A. (2010). The real effects of political uncertainty: Elections and investment sensitivity to stock prices. *Working Paper*.

Eesley, C. E. and Lee, Y. S. (2021). Do university entrepreneurship programs promote entrepreneurship? *Strategic Management Journal*, 42(4):833–861.

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

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

Ewens, M., Nanda, R., and Rhodes-Kropf, M. (2018). Cost of experimentation and the evolution of venture capital. *Journal of Financial Economics*, 128(3):422–442.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Frank, C., Sink, C., Mynatt, L., Rogers, R., and Rappazzo, A. (1996). Surviving the valley of death: A comparative analysis. *Journal of Technology Transfer*, 21(1-2):61–69.

Fuss, S., Szolgayova, J., Obersteiner, M., and Gusti, M. (2008). Investment under Market and Climate Policy Uncertainty. *Applied Energy*, 85(8):708–721.

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

Gallaugher, J. M. and Wang, Y.-M. (2002). Understanding Network Effects in Software Markets: Evidence from Web Server Pricing. *MIS Quarterly*, 26(4):303–327.

Gawel, E. and Lehmann, P. (2019). Should renewable energy policy be 'renewable'? *Oxford Review of Economic Policy*, 35(2):218–243.

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

Gentzkow, M. and Shapiro, J. M. (2006). Media bias and reputation. *Journal of political Economy*, 114(2):280–316.

Gentzkow, M. and Shapiro, J. M. (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1):35–71.

Ghanem, D. and Smith, A. (2020). What are the benefits of high-frequency data for fixed effects panel models? *Journal of the Association of Environmental and Resource Economists*, pages 710–968.

Goldenberg, S. (2010). Fate of US climate change bill in doubt after Scott Brown's Senate win.

Goldstein, A., Doblinger, C., Baker, E., and Anadón, L. D. (2020). Patenting and business outcomes for cleantech startups funded by the Advanced Research Projects Agency-Energy. *Nature Energy*, 5(10):803–810.

Gompers, P. and Lerner, J. (2001). The venture capital revolution. *Journal of Economic Perspectives*, 15(2):145–168.

Gompers, P. A., Gornall, W., Kaplan, S. N., and Strebulaev, I. A. (2020). How do venture capitalists make decisions? *Journal of Financial Economics*.

Gonzalez-Uribe, J. and Leatherbee, M. (2017). The Effects of Business Accelerators on Venture Performance: Evidence from Start-Up Chile. *The Review of Financial Studies*.

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

Guien, H. and Ion, M. (2016). Policy Uncertainty and Corporate Investment. *The Review of Financial Studies*, 29(3):523–564.

Hall, B. H. and Lerner, J. (2010). The financing of R&D and innovation. In *Handbook of the Economics of Innovation*, volume 1, pages 609–639. Elsevier.

Hallen, B. L., Cohen, S. L., and Bingham, C. B. (2020). Do accelerators work? If so, how? *Organization Science*, 31(2):378–414.

Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics*, 95(2):347–361.

Hamilton, J. T. (1995). Pollution as news: Media and stock market reactions to the Toxics Release Inventory data. *Journal of Environmental Economics and Management*, 28(1):98–113.

Hansen, S. and McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99:S114–S133.

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and Deliberation within the FOMC: A Computational Linguistics Approach. *The Quarterly Journal of Economics*, 133(2):801–870.

Hartley, P. R. and Medlock, K. B. (2017). The valley of death for new energy technologies. *Energy Journal*, 38(3):33–61.

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

Hildebrand, T., Puri, M., and Rocholl, J. (2017). Adverse Incentives in Crowdfunding. *Management Science*, 63(3):587–608.

Hochberg, Y. V. (2016). Accelerating entrepreneurs and ecosystems: The seed accelerator model. *Innovation Policy and the Economy*, 16(1):25–51.

Holtz-Eakin, D., Joulfaian, D., and Rosen, H. S. (1994). Sticking it Out: Entrepreneurial Survival and Liquidity Constraints. *Journal of Political Economy*, 102(1):53–75.

Howell, S. T. (2017). Financing Innovation : Evidence from R&D Grants. *American Economic Review*, 107(4):1136–1164.

Howell, S. T. (2020). Reducing information frictions in venture capital: The role of new venture competitions. *Journal of Financial Economics*, 136(3):676–694.

Howell, S. T. (2021). Learning from Feedback: Evidence from New Ventures. *Review of Finance*, 25(3):595–627.

Hudson, J. and Khazragui, H. F. (2013). Into the valley of death: Research to innovation. *Drug Discovery Today*, 18(13-14):610–613.

Huse, C. and Koptyug, N. (2021). Salience and policy instruments: Evidence from the auto market. *Journal of the Association of Environmental and Resource Economists*, pages 716–878.

Jacobsen, G. D. (2011). The Al Gore effect: An Inconvenient Truth and voluntary carbon offsets. *Journal of Environmental Economics and Management*, 61(1):67–78.

Johnstone, N., Hašcic, I., and Kalamova, M. (2010). Environmental Policy Design Characteristics and Technological Innovation.

Julio, B. and Yook, Y. (2012). Political uncertainty and corporate investment cycles. *Journal of Finance*, 67(1):45–84.

Kahn, M. E. and Kotchen, M. J. (2011). Business cycle effects on concern about climate change: The chilling effect of recession. *Climate Change Economics*, 2(03):257–273.

Kalamova, M., Johnstone, N., and Haščič, I. (2013). Implications of Policy Uncertainty for Innovation in Environmental Technologies: The Case of Public R&D Budgets. In Costantini, V. and Mazzanti, M., editors, *The Dynamics of Environmental and Economic Systems: Innovation, Environmental Policy and Competitiveness*, pages 99–116. Springer Netherlands, Dordrecht.

Kamien, M. I. and Schwartz, N. L. (1975). Market Structure and Innovation : A Survey. *Journal of Economic Literature*, 13(1):1–37.

Kaplan, S. N. and Lerner, J. (2010). It ain't broke: The past, present, and future of venture capital. *Journal of Applied Corporate Finance*, 22(2):36–47.

Kaplan, S. N. and Schoar, A. (2005). Private equity performance: Returns, persistence, and capital flows. *Journal of Finance*, 60(4):1791–1823.

Katz, M. L. and Shapiro, C. (1992). Product Introduction with Network Externalities. *The Journal of Industrial Economics*, 40(1):55–83.

Keller, T. R., Hase, V., Thaker, J., Mahl, D., and Schäfer, M. S. (2020). News media coverage of climate change in India 1997–2016: Using automated content analysis to assess themes and topics. *Environmental Communication*, 14(2).

Kerr, W. R., Lerner, J., and Schoar, A. (2011). The Consequences of Entrepreneurial Finance : Evidence from Angel Financings. *The Review of Financial Studies*, (27(1)):20–55.

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

Koellinger, P., Minniti, M., and Schade, C. (2007). "I think I can, I think I can": Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology*, 28(4):502–527.

Kruse, T., Mohnen, M., and Sato, M. (2020). Are financial markets aligned with climate action? New evidence from the Paris Agreement.

Kumar, S., Managi, S., and Matsuda, A. (2012). Stock prices of clean energy firms, oil and carbon markets: A vector autoregressive analysis . *Energy Economics*, 34(1):215–226.

Kyritsis, E. and Serletis, A. (2019). Oil Prices and the Renewable Energy Sector. *The Energy Journal*, 40(01):337–364.

Kölbel, J. F., Leippold, M., Rillaerts, J., and Wang, Q. (2020). Ask BERT: How regulatory disclosure of transition and physical climate risks affects the CDS term structure.

Lemoine, D. (2017). Green Expectations: Current Effects of Anticipated Carbon Pricing. *Review of Economics and Statistics*, 99(3):499–513.

Lerner, J. and Nanda, R. (2020). Venture Capital's Role in Financing Innovation: What We Know and How Much We Still Need to Learn. *Journal of Economic Perspectives*, 34(3):237–261.

Lerner, J., Schoar, A., Sokolinski, S., and Wilson, K. (2018). The globalization of angel investments: Evidence across countries. *Journal of Financial Economics*, 127(1):1–20.

Lewis, M. (2018). *The Fifth Risk*. W. W. Norton & Company.

Lopez-Claros, A., Dahl, A. L., and Groff, M. (2020). *The Challenges of the 21st Century*, page 3–29. Cambridge University Press.

Loughran, T. and Mcdonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1):35–65.

Lyons, E. and Zhang, L. (2018). Who does (not) benefit from entrepreneurship programs? *Strategic Management Journal*, 39(1):85–112.

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

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714.

Mcdonald, R. and Siegel, D. (1986). The Value of Waiting to Invest. *Quarterly Journal of Economics*, 101(4):707–727.

McKenzie, D. (2017). Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition. *American Economic Review*, 107(8):2278–2307.

Metrick, A. and Yasuda, A. (2021). *Venture capital and the finance of innovation.* John Wiley & Sons.

Minniti, M. and Bygrave, W. (2001). A dynamic model of entrepreneurial learning. *Entrepreneurship theory and practice*, 25(3):5–16.

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of Business Venturing*, 29(1):1–16.

Mowery, D. C., Nelson, R. R., and Martin, B. R. (2010). Technology policy and global warming: Why new policy models are needed (or why putting new wine in old bottles won't work). *Research Policy*, 39(8):1011–1023.

Muehlenbachs, L., Sinha, E. N., and Sinha, N. R. (2011). Strategic release of news at the EPA. *RFF Discussion Paper*.

Mukanjari, S. and Sterner, T. (2018). Do markets trump politics? Evidence from fossil market reactions to the Paris Agreement and the U.S. election.

Mullainathan, S. and Shleifer, A. (2005). The market for news. *The American Economic Review*, 95(4):46.

Nanda, R. and Kropf, M. R. (2016). Financing entrepreneurial experimentation. *Innovation Policy and the Economy*, 16(1):1–23.

Nanda, R. and Rhodes-Kropf, M. (2013). Investment cycles and startup innovation. *Journal of Financial Economics*, 110(2):403–418.

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

Neher, D. V. (1999). Staged financing: An agency perspective. *Review of Economic Studies*, 66(2):255–274.

Nemet, G. F., Zipperer, V., and Kraus, M. (2018). The valley of death, the technology pork barrel, and public support for large demonstration projects. *Energy Policy*, 119(October 2017):154–167.

Noailly, J., Nowzohour, L., and van den Heuvel, M. (2021). Heard the News ? Environmental Policy and Clean Investments. *CIES Research Paper*, (70).

Noailly, J., Nowzohour, L., and van den Heuvel, M. (2022). News coverage of US environmental policy: An analysis of major themes using topic modeling.

Noailly, J. and Smeets, R. (2015). Directing technical change from fossil-fuel to renewable energy innovation: An application using firm-level patent data. *Journal of Environmental Economics and Management*, 72:15–37.

Ozmel, U., Robinson, D. T., and Stuart, T. E. (2013). Strategic alliances, venture capital, and exit decisions in early stage high-tech firms. *Journal of Financial Economics*, 107(3):655–670.

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

Pástor, and Veronesi, P. (2013). Political uncertainty and risk premia. *Journal of Financial Economics*, 110(3):520–545.

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

Popp, D. (2020). Promoting Clean Energy Innovation at the State and Local Level. *Agricultural and Resource Economics Review*, 49(2):360–373.

Popp, D., Pless, J., Haščič, I., and Johnstone, N. (2020). Innovation and Entrepreneurship in the Energy Sector. *NBER Working Paper Series*.

Reboredo, J. C. and Ugolini, A. (2018). The impact of Twitter sentiment on renewable energy stocks. *Energy Economics*, 76:153–169.

Robb, A. M. and Robinson, D. T. (2014). The capital structure decisions of new firms. *Review of Financial Studies*, 27(1):153–179.

Ruhnka, J. C., Feldman, H. D., and Dean, T. J. (1992). The "living dead" phenomenon in venture capital investments. *Journal of Business Venturing*, 7(2):137–155.

Sada, D. and Muro, M. (2017). Cleantech venture capital: Continued declines and narrow geography limit prospects. Technical report, Brookings Institution Report.

Sadorsky, P. (2012). Correlations and volatility spillovers between oil prices and the stock prices of clean energy and technology companies. *Energy Economics*, 34(1):248–255.

Saha, D. and Muro, M. (2017). Cleantech venture capital : Continued declines and narrow geography limit prospects. *Brookings Institution Report*, pages 1–14.

Sato, M., Kruse, T., Burke, J., Mohnen, M., and Pope, P. (2020). Does it pay for firms to go green?

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

Scott, E. L., Shu, P., and Lubynsky, R. M. (2020). Entrepreneurial uncertainty and expert evaluation: An empirical analysis. *Management Science*, 66(3):1278–1299.

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

Shapiro, J. M. (2016). Special interests and the media: Theory and an application to climate change. *Journal of Public Economics*, 144:91–108.

Smith, B. and Viceisza, A. (2018). Bite me! ABC's Shark Tank as a path to entrepreneurship. *Small Business Economics*, 50(3):463–479.

Song, Y., Ji, Q., Du, Y.-J., and Geng, J.-B. (2019). The dynamic dependence of fossil energy, investor sentiment and renewable energy stock markets. *Energy Economics*, 84:104–564.

Stokey, N. L. (2016). Wait-and-see: Investment options under policy uncertainty. *Review of Economic Dynamics*, 21:246–265.

TheEconomist (2021). Billions are pouring into the business of decarbonisation. *The Economist*.

Tian, X. (2011). The causes and consequences of venture capital stage financing. *Journal of Financial Economics*, 101(1):132–159.

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

Timmons, J. A. and Bygrave, W. D. (1986). Venture capital's role in financing innovation for economic growth. *Journal of Business Venturing*, 1(2):161–176.

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

Venture Kick (2019). Annual Report. pages 1–28.

Viscusi, W. K. (1983). Frameworks for Analyzing the Effects of Risk and Environmental Regulations on Productivity. *The American Economic Review*, 73(4):793–801.

Weyant, J. P. (2011). Accelerating the development and diffusion of new energy technologies: Beyond the "valley of death". *Energy Economics*, 33(4):674–682.

Yu, S. (2019). How Do Accelerators Impact the Performance of High-Technology Ventures? *Management Science*.

Appendix

A Appendix for Chapter 1

Venture Kick's grading categories in detail

Table A.1: Grades given in each stage

	Stage 1	Stage 2	Stage 3
Grade 1	Visionary & Leadership	Visionary & "Winning"	Visionary & Leadership
Grade 2	"Winning" & Communication	Team & Complementary	"Winning" & Communication
Grade 3	Credibility & Expertise	Value Prop. & Opportunity	Strong team & Skills
Grade 4	Team & Complementary	USP & IP	USP & Business opportunity
Grade 5	Network & References	Business Model	BP & Roadmap and Financing
Average	People	People & Project	People & Project
Grade 6	Value prop. & Opportunity	Market & Customers	Sold products & Projects
Grade 7	Market & Customers	Experts & Advisory	Partnerships & Alliances
Grade 8	USP & IP	Summary & Investor attraction	Marketing & "Buzz"
Grade 9	Business Model	Roadmap & Process	Network & Sources
Grade 10	Achievements & Roadmap	Financial plan & Strategy	Fund raised & Incorporation
Average	Project	Achievements & Development	Achievements & Development
Grade 11	"Gut Feeling"	"Gut Feeling"	"Gut Feeling"

Venture Kick's publicity

Venture Kick is a respected institution

Venture Kick's publicity is valuable because it originates from a trusted institution. These are endorsements from a member of Switzerland's largest private science and innovation foundation and a member of Swisscom, Switzerland's biggest telecom company :

Dr. Pascale Vonmont, Gebert Rüf Stiftung

Venture Kick closes a key gap in the innovation chain by early identification and promotion of high-risk, highpotential business ideas in order to create the markets of the future. The highly respected promotion channels guarantee national and international visibility for the startups.

Roger Wüthrich-Hasenböhler, Swisscom

All successful startups in Switzerland have come out of the tough Venture Kick process. Venture Kick is the most important startup initiative and we're proud to support this unique private program.

Source: Venture Kick's website

The case of Imverse: a startup that won stage 1 in 2017

Venture Kick's article

Four new breakthrough startup projects win CHF 10'000 with Venture Kick

09.02.2017

Four projects won the first stage of Venture Kick. They are getting ready to present on the second stage of the program, where they might win another CHF 20'000. If they make it at the final, they will get CHF 130'000 in total.

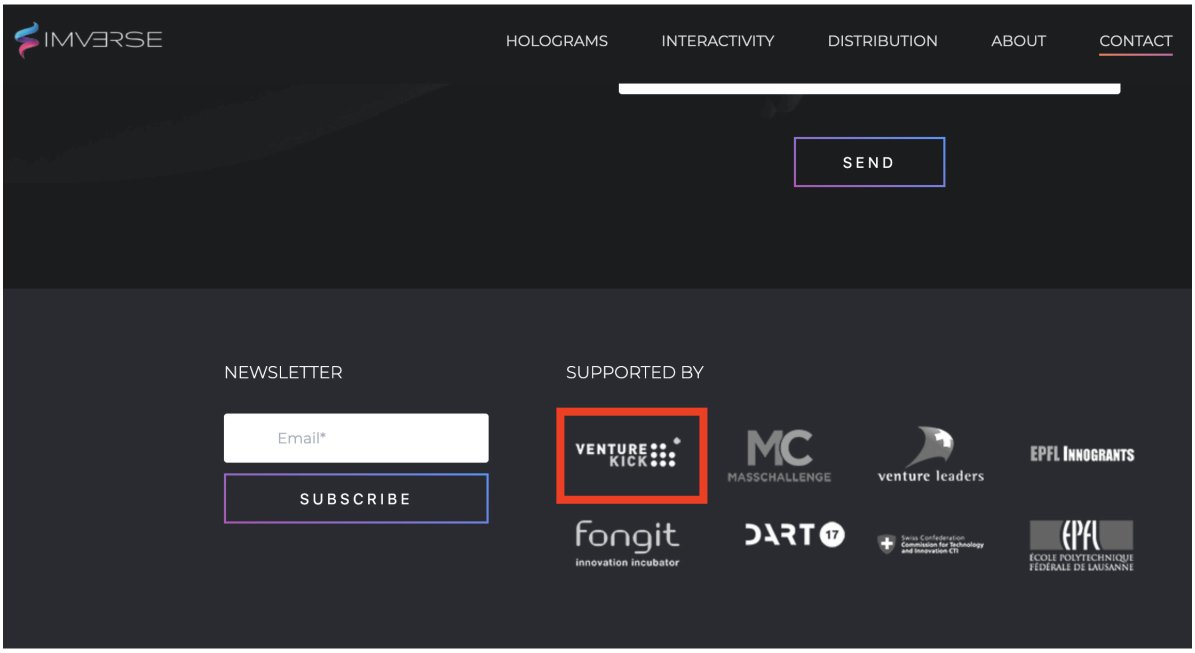
Imverse (EPFL) – Software, VR/AR: The young startup aims at democratizing virtual and augmented reality creation. So far, only highly skilled technical professionals and big budgets can create immersive experiences. To address this problem, Imverse plans to offer software licenses and B2B services for anyone to create their first mixed reality content: room-scale, realistic, fast, inexpensive and personalized.

Source: Venture Kick's website

Venture Kick's tweet

Source: Twitter

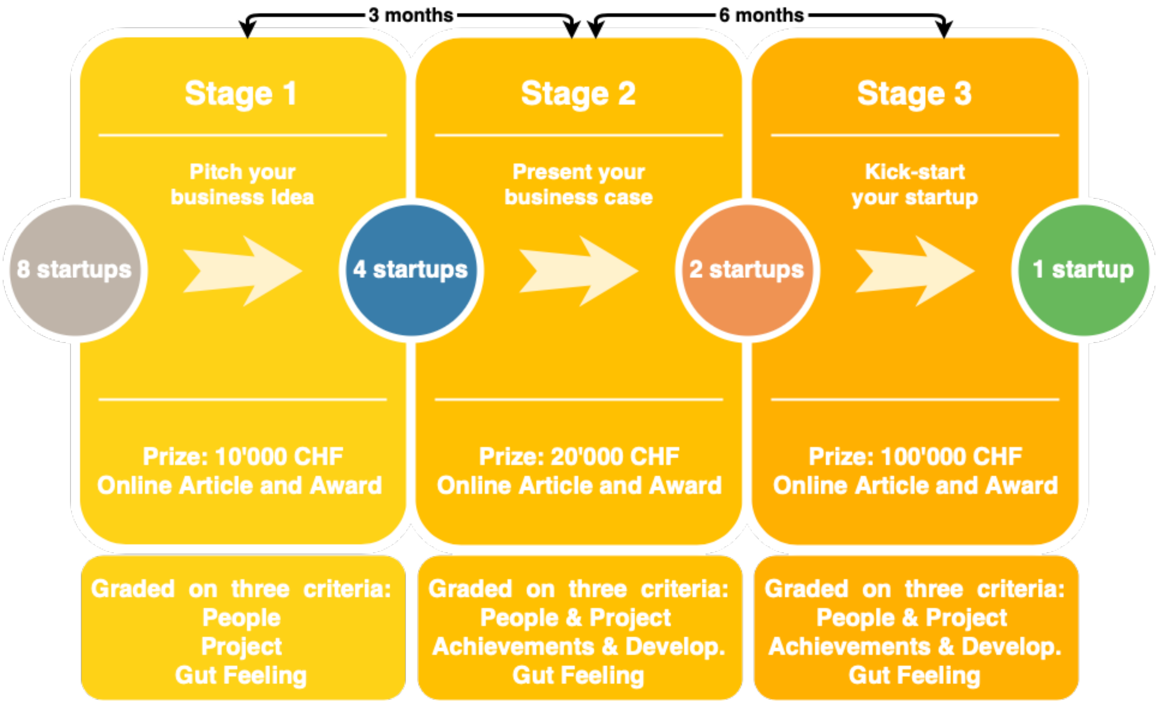
Venture Kick's logo on Imverse's website



Source: Imverse's website

The Venture Kick process

Figure A.1: The three stages of Venture Kick



Source: EPFL

Principal component analysis using our eleven grades, in stage 1

Table A.2: Eigenvalues of the components of our PCA, Stage 1

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp 1	8.327	7.487	0.757	0.757
Comp 2	0.841	0.397	0.076	0.834
Comp 3	0.444	0.142	0.040	0.874
Comp 4	0.301	0.046	0.027	0.901
Comp 5	0.255	0.021	0.023	0.924
Comp 6	0.235	0.060	0.021	0.946
Comp 7	0.175	0.034	0.016	0.962
Comp 8	0.141	0.029	0.013	0.974
Comp 9	0.112	0.009	0.010	0.985
Comp 10	0.103	0.036	0.009	0.994
Comp 11	0.067	0.000	0.006	1.000

Table A.3: Component loadings for the first four components, Stage 1

Grades	Comp 1	Comp 2	Comp 3	Comp 4	Unexplained
Visionary & Leadership	0.298	0.402	0.329	0.066	0.078
Winning & Communication	0.269	0.587	0.299	0.019	0.069
Credibility & Expertise	0.311	-0.081	-0.325	0.084	0.141
Team & Complementarity	0.294	0.278	-0.311	0.290	0.147
Network & References	0.303	-0.020	-0.579	0.194	0.076
Value proposition & Opp.	0.286	-0.359	0.403	0.477	0.071
Market & Customers	0.316	-0.022	-0.153	-0.017	0.159
USP & IP	0.283	-0.456	0.121	-0.114	0.150
Business Model	0.316	-0.269	0.219	-0.059	0.085
Achievements & Roadmap	0.303	0.009	-0.061	-0.780	0.052
Gut Feeling	0.334	-0.023	0.112	-0.117	0.060

Estimating whether certification impacts winners or losers more strongly.

We can test the claim that the impact of Venture Kick on survival is concentrated on losing startups by comparing the survival rates of startups in our sample with the typical survival rates of young Swiss or American ventures. The Swiss Federal Statistical Office (FSO) and the U.S. Bureau of Labor Statistics (BLS) provide these figures. They calculated that, in the past decade, the survival rate of startups two years after birth has been around 67–70 percent, 60 percent after three years, then 54 percent and finally 50 percent after five years.[28] For Venture Kick's participants, these numbers are respectively 61, 58, 55 and 52 percent. In other words, in the short term, startups that participate in the competition have lower survival rates than typical young companies but then after four years they start having slightly higher survival rates than typical startups. Additionally, we use the model in Table 1.3 to predict the survival rate of a startup of average quality as a function of whether it wins or loses stage 1. If this startup wins, the predicted survival rate after 2 years is 0.69, very close to the FSO value for the average young Swiss company. The predicted survival rate if it loses reaches 0.48, more than 22 percentage points below the average survival rate in Switzerland. Taken together, these elements provide evidence that losers will cut their losses much quicker than they would have otherwise, thus avoiding costly wastage of resources. Given that this effect of Venture Kick's certification of quality on survival rates seems to be focused on the losers, it makes sense that it is only the performance in stage 1 that causally impacts survival. Indeed, losing in stages 2 and 3 will only happen after having won stage 1 and received praise for their quality. Losing in these later stages should thus not cause entrepreneurs to drastically rethink their probability of success.

[28] These statistics are available on the FSO's website: https://www.bfs.admin.ch/bfs/.../taux-de-survie.assetdetail.10687119.html and the BLS's: https://www.bls.gov/bdm/us_age_naics_00_table7.txt

Alternative identification strategies for main effect

Table A.4: Control variables: Effect of winning stage 1

	Survival after the competition			Funding after comp.	
	(1)	(2)	(3)	(4)	(5)
	Survival 2	Survival 2	Survival 2	Rounds	Funding
Wins Stage 1	0.341***	0.399***	0.410***	0.582***	0.148***
	(0.0904)	(0.0894)	(0.0889)	(0.184)	(0.0553)
Wins Stage 1 x ICT	-0.310**	-0.311**	-0.311**	-0.582***	-0.153**
	(0.123)	(0.125)	(0.127)	(0.219)	(0.0716)
Standardized grade	0.122***	0.143***		0.403***	0.134***
	(0.0405)	(0.0409)		(0.0937)	(0.0292)
Grade x ICT	0.00218	-0.0149		-0.139	-0.0692**
	(0.0569)	(0.0587)		(0.118)	(0.0331)
Motivation	0.202***				0.170***
	(0.0482)				(0.0486)
Average Grade			0.133***		
			(0.0413)		
Grade x ICT			-0.0201		
			(0.0594)		
Order FE	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Competition FE	Yes	Yes	Yes	Yes	Yes
University FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Observations	671	671	671	909	909
R^2	0.396	0.392	0.377	0.320	0.355

Note: This table presents estimates of the effect of winning stage 1 using equation (1.2) with additional controls. Columns (1) to (3) focus on the survival outcome. Column (1) controls for motivation, Column (2) for the order of passage during the competition and Columns (3) uses an alternative grade variable. Columns (4) and (5) introduce a new dependent variable the number of funding rounds secured and test our control variables on *Funding* The level of observation is a startup in a competition round. Errors are clustered by competition round. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

Table A.5: Control variables: Effect of winning stage 3

	Survival after the competition			Funding after comp.	
	(1)	(2)	(3)	(4)	(5)
	Survival 2	Survival 2	Survival 2	Rounds	Funding
Wins Stage 3	-0.0127	-0.0135	-0.0310	0.233	0.143
	(0.0517)	(0.0483)	(0.0533)	(0.771)	(0.165)
Wins Stage 3 x ICT	0.253*	0.210*	0.239*	0.317	0.113
	(0.127)	(0.119)	(0.126)	(1.149)	(0.240)
Standardized grade	0.0572	0.0580		0.427	0.0497
	(0.0552)	(0.0544)		(0.372)	(0.102)
Grade x ICT	0.0467	0.0485		-0.0362	0.116
	(0.0706)	(0.0733)		(0.531)	(0.118)
Motivation	-0.0337				-0.00756
	(0.0851)				(0.0935)
Average Grade			0.0622		
			(0.0561)		
Grade x ICT			0.0501		
			(0.0769)		
Order FE	No	Yes	No	No	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Competition FE	Yes	Yes	Yes	Yes	Yes
University FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Observations	178	178	178	216	216
R^2	0.468	0.475	0.472	0.457	0.518

Note: This table presents estimates of the effect of winning stage 3 using equation (1.2) with additional controls. Columns (1) to (3) focus on the survival outcome. Column (1) controls for motivation, Column (2) for the order of passage during the competition and Columns (3) uses an alternative grade variable. Columns (4) and (5) introduce a new dependent variable the number of funding rounds secured and test our control variables on *Funding* The level of observation is a startup in a competition round. Errors are clustered by competition round. *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

Table A.6: Errors-in-variables regression: Effect of winning stages 1, 2 and 3

	Baseline		EiV
	(1)	(2)	(3)
	Survival 2	Survival 2	Survival 2
Stage 1			
Wins Stage 1	0.220***	0.354***	0.257***
	(0.0683)	(0.0456)	(0.0811)
Standardized grade	0.119***	0.0610***	0.132***
	(0.0353)	(0.0222)	(0.0511)
Stage 2			
Wins Stage 2	-0.0541	0.101**	0.0608
	(0.0666)	(0.0467)	(0.0574)
Standardized grade	0.148***	0.0427*	0.0768*
	(0.0429)	(0.0224)	(0.0407)
Stage 3			
Wins Stage 3	0.0477	0.0782**	0.0569
	(0.0479)	(0.0390)	(0.0371)
Standardized grade	0.0578	0.0226	0.0407
	(0.0413)	(0.0190)	(0.0298)
Year FE	Yes	No	No
Competition FE	Yes	No	No
Industry FE	Yes	No	No
University FE	Yes	No	No
Observations stage 1	671	671	671
Observations stage 2	353	353	353
Observations stage 3	178	178	178

Note: Table presents results of standard as well as errors-in-variables regressions. Column (1) reports our baseline specification, Column (2) a standard regression without fixed effects which serves as a benchmark for Column (3), which shows our errors-in-variables regression with our grade variable having a reliability of 0.7. The level of observation is a startup in a competition round. Errors are clustered by competition round in columns (1). *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

Table A.7: Effect of winning stages 1 and 3 using an errors-in-variables regression

	Baseline		EiV
	(1) Survival 2	(2) Survival 2	(3) Survival 2
Stage 1			
Wins Stage 1	0.410***	0.499***	0.512***
	(0.0870)	(0.0601)	(0.0959)
Wins Stage 1 x ICT	-0.315**	-0.323***	-0.354*
	(0.123)	(0.0905)	(0.210)
Standardized grade	0.137***	0.0558*	0.0459
	(0.0404)	(0.0295)	(0.0617)
Grade x ICT	-0.0155	0.0107	0.0333
	(0.0571)	(0.0445)	(0.143)
Stage 3			
Wins Stage 3	-0.0205	0.00316	0.00697
	(0.0506)	(0.0479)	(0.0271)
Wins Stage 3 x ICT	0.250*	0.207**	0.200**
	(0.125)	(0.0806)	(0.0793)
Standardized grade	0.0562	0.0244	0.0216
	(0.0550)	(0.0236)	(0.0282)
Grade x ICT	0.0470	0.0125	0.0199
	(0.0727)	(0.0380)	(0.0648)
Year FE	Yes	No	No
Competition FE	Yes	No	No
University FE	Yes	No	No
Industry FE	Yes	Yes	Yes
Observations Stage 1	671	671	671
Observations Stage 3	178	178	178

Note: This Table presents results of standard as well as errors-in-variables regressions. The top panel shows the effect of winning stage 1 and the bottom panel the effect of winning stage 3. Columns (1) reports our baseline specification. Columns (2), a standard without fixed effects, serves as a benchmark for Column (3), that shows our errors-in-variables regressions where the ICT grade interaction has a reliability of 0.8. The grades attributed to science startups are assumed to be reliable. The level of observation is a startup in a competition round. Errors are clustered by competition round in column (1). *, **, and *** denote statistical significance at 10%, 5%, and 1% levels, respectively.

Propensity Score Matching, balance across treatment and comparison groups

Figure A.2: Propensity score balance across treatment and comparison groups, matching on grade, industry, university, gender and year, Stage 1

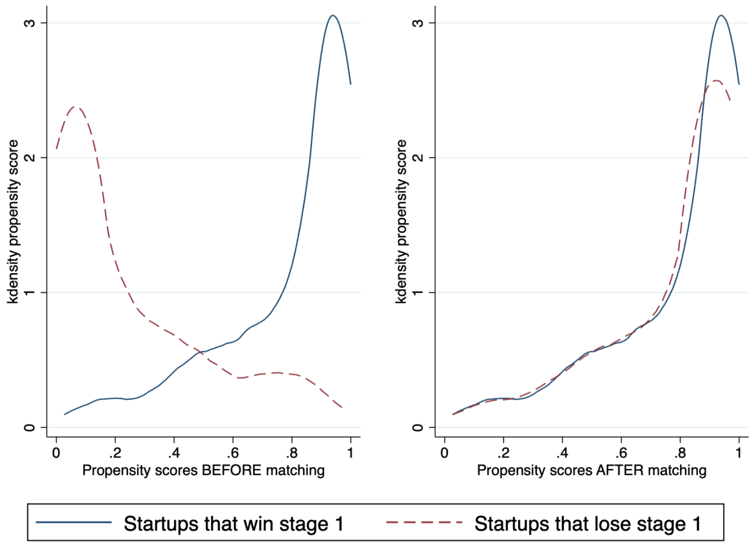
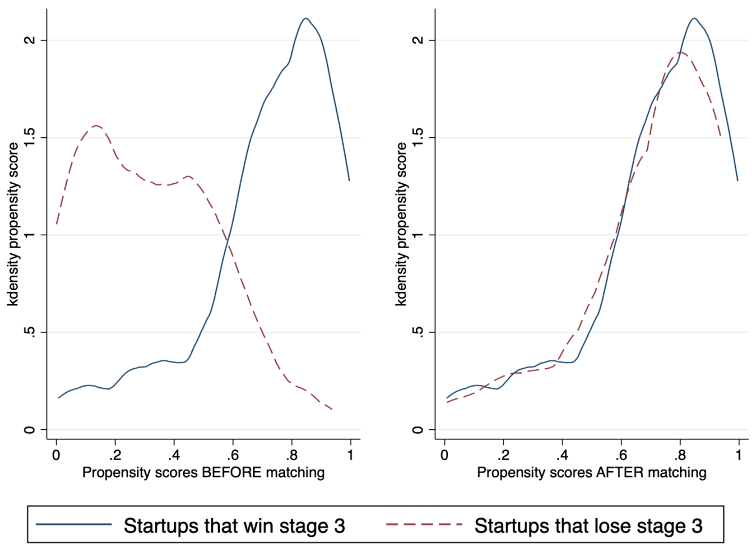


Figure A.3: Propensity score balance across treatment and comparison groups, matching on grade, industry, university, gender and year, Stage 3



B Appendix for Chapter 2

There is no additional material for this chapter.

C Appendix for Chapter 3

Additional tables and figures: EnvP index

Table C.1: Newspaper list

Newspaper	Available since	% in Dow Jones	% in Query
New York Times	1 June 1980	22.5%	21.9%
Washington Post	6 January 1982	15.3%	17.9%
Wall Street Journal	13 June 1979	9.8%	12.8%
Tampa Bay Times	11 June 1986	11.5%	12.3%
Houston Chronicle	2 February 1985	13.8%	11.1%
Dallas Morning News	18 January 1984	10.8%	7.5%
San Francisco Chronicle	4 January 1985	6.2%	7.4%
San Jose Mercury News	2 January 1994	3.4%	3.9%
Boston Herald	26 July 1991	5.0%	2.7%
San Diego Union Tribune	31 December 2010	1.7%	2.4%

Figure C.1: EnvP - WSJ versus Engle et. al

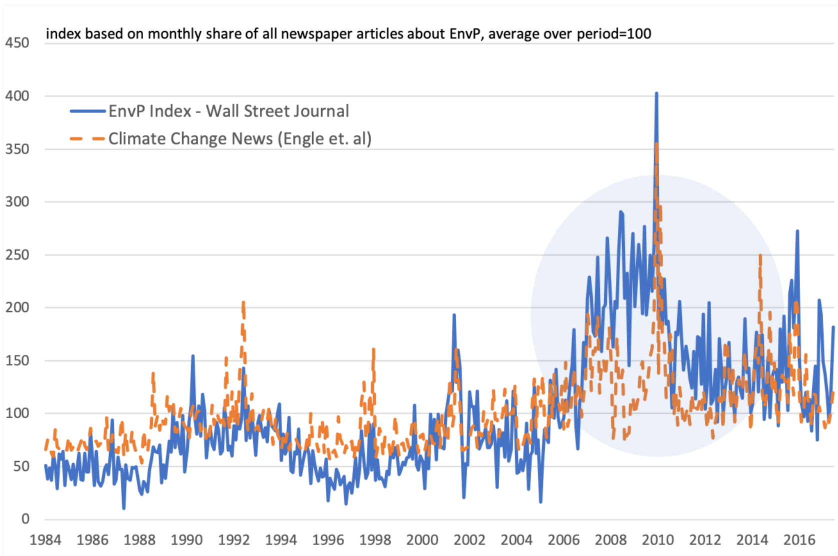
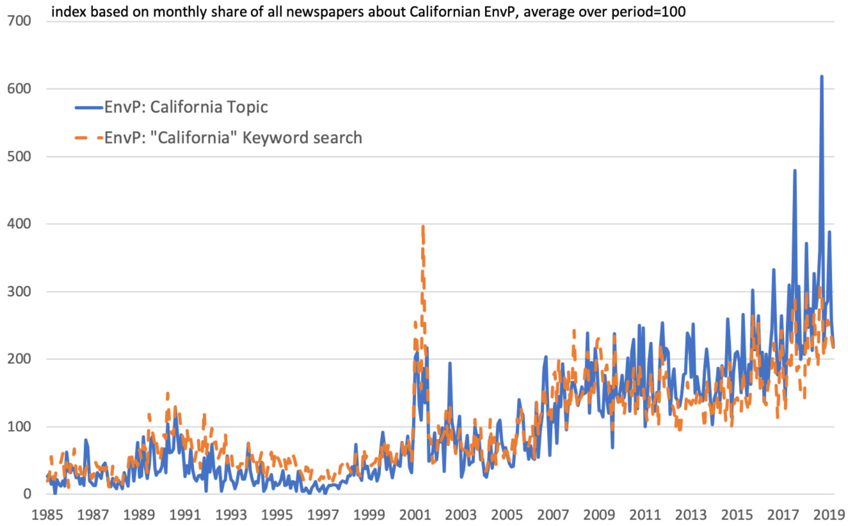
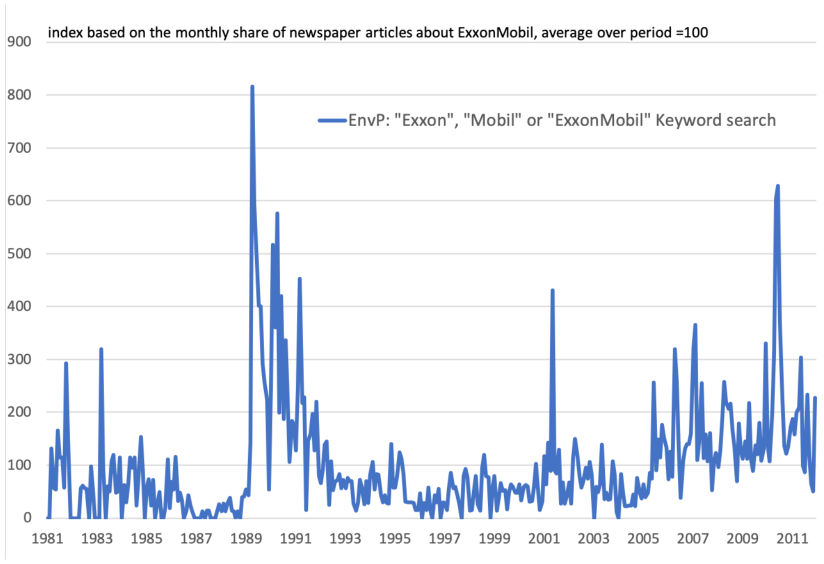


Figure C.2: California EnvP indices



The figure plots two alternative versions of a sub-EnvP index for the state of California: 1) a version in which we select articles from the 80,045 articles composing our EnvP index using the "California" keyword, 2) a version in which we plot the index resulting from our California topic in the topic modeling exercise. Both approaches capture similar trends.

Figure C.3: ExxonMobil index



The figure plots the evolution of a sub-EnvP index capturing discussions around ExxonMobil in environmental policy news, using the combination of keywords 'Exxon', 'Mobil' or 'ExxonMobil' into our set of EnvP articles. This index correctly identifies the main events around 1) the Exxon Valdez oil spill in March 1989, which led to the implementation of the Oil Pollution Act in the following year, 2) the US withdrawal from the Kyoto protocol and debates around a letter from Exxon to Bush to outcast IPCC's chairman in February/March 2001, 3) Exxon's attacks on the Clean Air Act in 2006, 4) the request from Exxon's shareholders to disclose potential economic impacts of the Paris accord on the company in May 2017.

The Support Vector Machines (SVM) Algorithm

Given a training set $T = ((x_j, y_i), j = 1..n)$ where x_j are the input variables (i.e. the features and their tf-idf score in each article) and y_i is the corresponding output value (i.e. the assigned label for each article), SVM methods will fit a model to this training set, for a given set of parameters. The algorithm builds on the idea that the numerical representation of each text/article are data points in a multivariate space of features. Based on the word content of each article the algorithm aims to find two hyperplanes separating the two classes of data (i.e., environmental policy articles and the rest). The classifier maximizes the distance between these hyperplanes, also called the margin. Articles or vectors that lie on one of these hyperplanes, i.e. particularly ambiguous articles that were hard to classify, are called support vectors. The decision boundary, that separates relevant from irrelevant articles is the hyperplane that lies at the mid-point of the margin.

Choosing the optimal hyperparameters

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

Evaluating the classifier's performance

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

Topic modeling

LDA is a generative statistical model of a corpus made of D documents — newspaper articles — and V unique terms. This topic model estimates K topics each of which is a distribution $\beta_k \in \triangle^V$ over all the unique terms V present in our articles.

The pre-processing

We use the pre-processed corpus of 80,045 articles identified in Section 3.2 and do some additional noise filtering. We tested the robustness of our preprocessing steps with the preText package in R (Denny and Spirling, 2018). The package computes a preText score for a range of text preprocessing specifications indicating whether any given specification is likely to yield 'unusual' results with respect to alternative ways to preprocess. We find that none of our initial preprocessing steps is significantly at risk of leading to unusual results. We include mono-, bi- and trigrams and use a tf-idf approach to filter out words that are either too rare or too common.

Selecting the optimal number of topics

The choice of the number of topics in the model is a critical step. Choosing too low a number tends to result in broad topics that miss the less prevalent but nonetheless important subtopics whereas choosing too high a number can lead to topics that are excessively narrow.

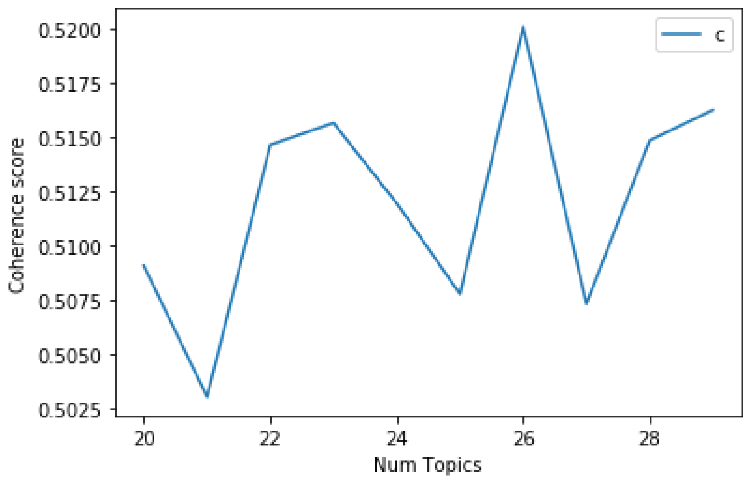
First, we compute the coherence score, a measure indicating how semantically similar the top features of any given topic are to one another. Figure C.4 reports coherence scores for LDA models with K ranging from 20 to 30. We favor this intermediate number of topics because while a higher number of topics tends to improve the statistical goodness-of-fit, it also makes the topics less easy to interpret. Next, we inspect the different outputs of these models to determine which one has the clearest and most interesting topics. We settle on a model with $K = 26$, the maximum in Figure C.4.

This model precisely identifies multiple topics we are particularly interested in such as policy discussions related to renewable energy, automobile emissions regulations and international climate negotiations. However, out of the 26 topics, one has no interpretable meaning and as a result we discard it as noise and do not take it into consideration when we assign topics to our articles. Table C.2 shows our 25 topics ranked from most to least prevalent in our sample. A more detailed analysis of the various topics is provided in a companion paper (Noailly et al., 2022).

Table C.2: Topic interpretation and classification (ranked by size).

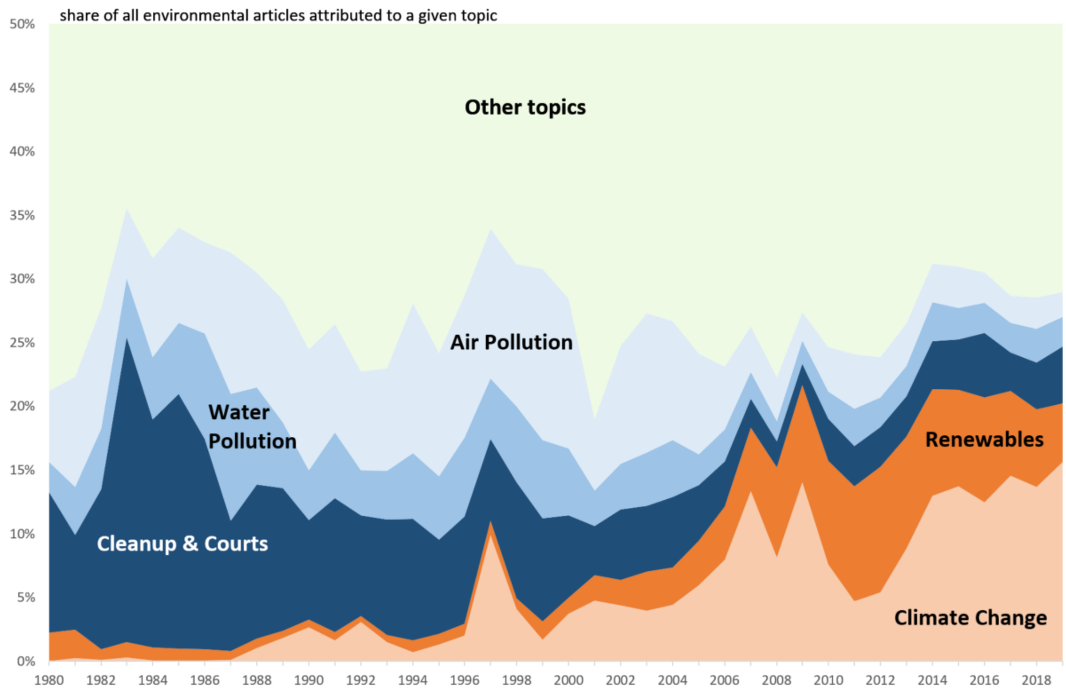
Topic	#	Topic	#	Topic	#
Climate Change	19	Oil & Gas production	15	Vehicle Fuels	12
EPA & Federal Gov.	5	Intl. Climate Negotiations	18	Waste & Recycling	26
Cleanups & Courts	17	Texas	11	Green Buildings	25
Government Budgets	3	Renewables	6	North-East Region	8
Air Pollution	9	Env. Conservation	4	Offshore Oil Drilling	7
Congress & Policy	13	Water Pollution	1	Nuclear Power	21
Businesses & Investments	22	Climate Science	16	Coal Industry	10
Presidents & Campaigns	23	California	14		
Power & Utilities	24	Automobile Industry	2		

Figure C.4: Topic coherence



To attribute topics to articles we use the fact that LDA models each article as a distribution over different topics so that each article d has an attribution to topic k, θ_d^k, in percentage. We could simply count all articles with $\theta_d^k > 0$ for any given topic. However, we aim to get rid of 'noisy articles', those whose attribution to a topic is below a critical threshold, $\theta_d^k < \theta_d^{k\min}$. Alternatively, we could simply choose to only pick one dominant topic per article. However, newspaper articles arguably talk about more than one relevant sub-topic of environmental policy and we aim to capture details of the U.S. policy debate. In addition, we believe that contemporary observers may draw information on policy issues, even if it is only mentioned as a side topic. To strike a middle ground, we use a cutoff α of 10%, meaning that only articles associated with any given topic with more than 10% probability, $\theta_d^k > 0.1 \ \forall \ d, k$ are counted.

Figure C.5: Evolution of media attention on environmental policy topics over time



The figure shows how media attention on these topics varies over time. In the 80s and 90s, the most important topics were 'Cleanup and Courts', 'Water Pollution' and 'Air Pollution'. More recently, 'Renewables' and 'Climate Change' have become central topics in the media.

Additional statistics and results: Venture Capital

Descriptive statistics - firm-level estimations and VAR

Table C.3 reports descriptive summary statistics for the main continuous variables used in our study of VC investments for the period between January 1998 and March 2019. The first panel shows the variables that are common to both the VAR and Panel analysis, taken from our monthly VAR dataset. The second panel displays statistics about the monthly number of VC deals in the United States that we use in our VAR analysis. Finally, the last panel reports the VC-related variables that we use in our panel analysis. The discrepancy between the number of observations reported in the last panel and in Table 3.3 comes from the fact that, in our analysis, we drop all the observations where the age variable is negative as it most likely indicates an error in the data. In other words we drop around 30,000 observations that were recorded before the startups' official founding date.

Table C.3: Summary statistics for venture capital analysis

	Obs.	Mean	Std. Dev	Min	Max
Environmental Policy Indices:					
EnvP Index	255	119.75	47.34	40.89	258.55
EnvP-RE Index	255	166.28	119.35	0.00	465.94
Economic Control Variables:					
YoY GDP Growth	243	2.19	1.75	-4.92	5.64
Oil price (WTI)	255	57.68	28.59	11.35	133.88
Fed Funds Rate	255	2.09	2.11	0.07	6.54
VAR: VC Variables					
Number of clean energy VC deals	255	15.58	11.25	0.00	48.00
Number of renewables VC deals	255	4.17	4.07	0.00	19.00
Panel: VC Variables					
Number of VC deals, per firm-quarter	1089760	0.06	0.24	0.00	3.00
VC amount raised (in mio), if funded	65061	12.16	36.71	0.00	3500.00
Age when funded (in years)	63989	5.18	4.40	0.00	33.50

Firm-level panel estimation: robustness analysis

VAR: robustness analysis

We provide further robustness analysis in Figure C.6. Our baseline response is very similar to the response from a bi-variate VAR model with EnvP-RE and our VC data, as well as the response from this bivariate model with reversed ordering. The responses from VAR models fitted on variables in levels and on a shorter sample from 2006 to 2019,[29] while potentially smaller especially in the longer-run, remain positive and significant.

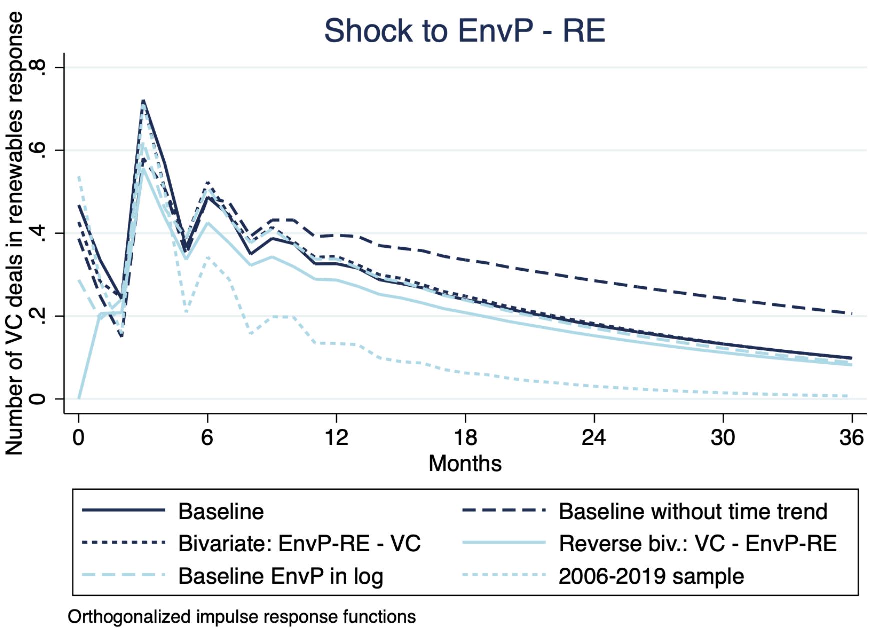
[29] 2006 is the start of the rapid rise in interest for renewable energy.

Table C.4: Probit model: Relationship between EnvP media coverage and VC investment in cleantech

	(1) Funded (Q+1)	(2) Funded (Q+1)	(3) Funded (Q+1)
Log EnvP index	0.0456***	-0.0564***	-0.103***
	(0.0107)	(0.0175)	(0.0177)
Cleantech	-1.466***	-1.331***	-1.601***
	(0.215)	(0.214)	(0.230)
Log EnvP x Cleantech	0.311***	0.288***	0.235***
	(0.0432)	(0.0456)	(0.0465)
Log Sentiment Index			-0.0637***
			(0.00446)
Log Sentiment x Cleantech			0.0889***
			(0.0214)
Industry FE	Yes	Yes	Yes
Other controls	Yes	Yes	Yes
Quarter FE	No	Yes	Yes
Year FE	No	Yes	Yes
State FE	No	Yes	Yes
Industry-time trend	No	Yes	Yes
Series FE	No	Yes	Yes
Observations	1056221	1056218	1056218
Firms	35637	35637	35637

Table presents results of a Probit regression. Other controls include age, oil price, a time trend, GDP and the fed fund rate. Standard errors are clustered at the company level.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure C.6: Renewable energy venture capital deals responses to EnvP-RE Shock, VAR fit to Monthly U.S. data under alternative specifications and samples



Additional statistics and results: Stock Returns

Descriptive statistics - firm-level estimations and VAR

We consider the following 5 market risk factors: 1) MKTR is the market factor measured as the difference between the returns of diversified portfolios of the overall market and the safe interest rate at the end of month t, 2) SMB is the size factor measured as the difference between the returns of diversified portfolios consisting of stocks of small firms and big firms at the end of month t, 3) HML is the value factor measured as the difference between the returns of diversified portfolios comprising stocks of firms with a high book-to-market equity ratio and firms with a low book-to-market equity ratio at the end of month t, 4) RMW is the profitability factor measured as the difference between the returns of diversified portfolios consisting of stocks with robust and weak profitability at the end of month t, 5) CMA is the investment factor measured as the difference between the returns of diversified portfolios consisting of stocks with low (conservative) and high (aggressive) investment.

Industry classifications are defined according to ICB as follows: Basic Materials, Consumer Discretionary, Consumer Staples, Energy, Health Care, Industrials, Technology, Telecommunications and Utilities. We exclude observations with negative equity or sales values and observations where growth in total assets was larger than 100% in absolute value.

Table C.5 shows summary statistics of all the variables used in the panel data analysis. In the estimations, environmental policy innovations and CO_2 emissions have been standardized to a mean of zero and a unit standard deviation and all financial variables are included in logs.

Table C.5: Firm-level panel estimations, stock returns - summary statistics

	Observations	Mean	Std. Dev	Min	Max
Environmental Policy Innovations:					
EnvP innovations	164	1.8	25	-57	82
EnvP net sentiment innovations	164	0.0023	0.0181	-0.0422	0.0518
Policy Exposure Variables:					
AVG Scope-1 Emissions (mln tCO2)	49143	5	17	0.0005	128
AVG Scope-1 Emission Intensity (tCO2/ market cap)	49143	564	1462	0.32	11581
Pre-sample Scope-1 Emissions (mln tCO2)	34079	7	21	0.000006	150
Financial Variables:					
Excess returns	49143	-0.08	0.15	-0.88	0.60
Leverage (debt/equity)	49143	1.5	7.1	0	347
Firm size (mln market cap)	49143	26.3	56.4	0.05	1167.2
Profitability (return on assets)	49143	8.1	5.2	0.03	60.8
Dividends per share	49143	1.2	1.8	0.009	33.1

The environmental policy innovations are residuals extracted from an AR(7) and AR(6), respectively. All financial variables are GDP deflated. The sample excludes the recession period associated with the GFC.

Table C.6 shows summary statistics of all the variables used in the VAR analysis.

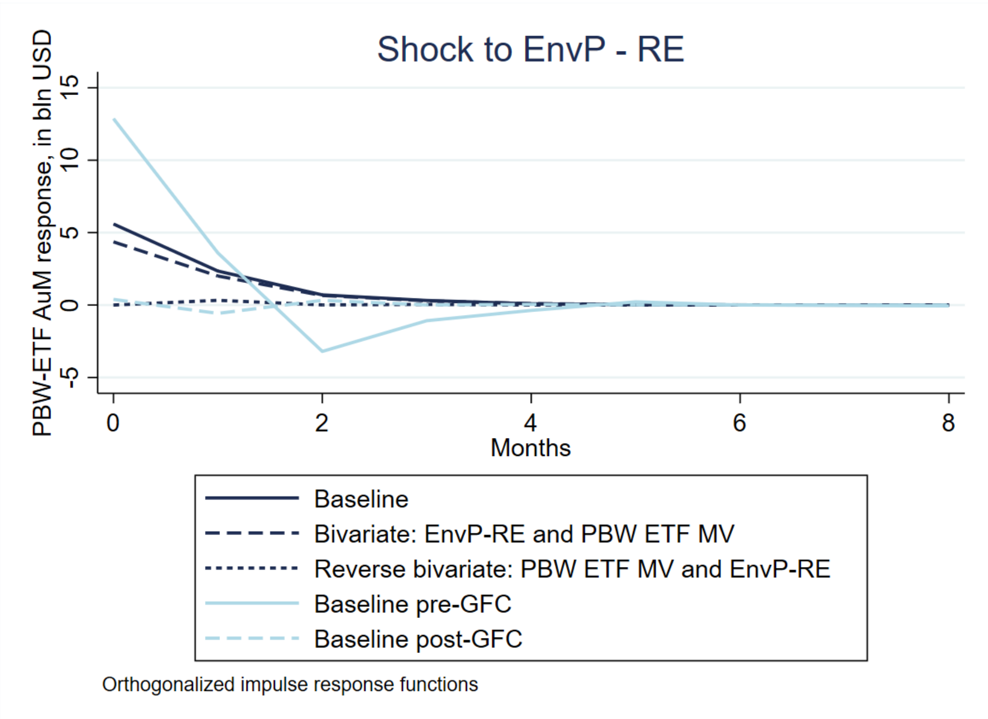
Table C.6: VAR models, ETF clean energy - summary statistics

	Observations	Mean	Std. Dev	Min	Max
Environmental Policy Index:					
Δ Log EnvP Index	149	0.00	0.32	-1.02	0.78
Economic Control Variables:					
Δ Log Oil (WTI)	149	0.00	0.08	-0.25	0.21
Δ Fed funds rate	149	0.03	0.08	-0.27	0.25
Δ Log NYSE Tech 100	149	0.01	0.03	-0.13	0.09
ETF Variables:					
Δ PBW ETF Assets under Management	149	4.07	38.94	-102.22	196.64

The sample excludes the recession period associated with the GFC as in our VAR.

VAR: robustness analysis

Figure C.7: PBW-ETF AuM responses to EnvP-RE Shock, VAR fit to Monthly U.S. data under alternative specifications and samples



D Appendix for Chapter 4

The newspapers in our index

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

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

A thesaurus-based index of environmental policy uncertainty.

We construct an extensive list of keywords based words extracted from a variety of sources such as the list of modal words in Tobback et al. (2018), our own set of manually labelled newspaper articles and other sources that led to a list of 270 unigrams and 160 bigrams. The main objective of this exercise is to investigate whether a more sophisticated keyword-based approach to identifying environmental policy uncertainty – arguably a fairer comparison to our more complex machine-learning algorithm – can do better than the 'uncertain*' approach commonly used in the literature. To do so, we collect a list of unigrams and bigrams which reflect our concept of environmental policy uncertainty. To mitigate overfitting, we do not select keywords based on their performance in the training set.[30] The query is made up of keywords shown in Table D.2.

Performance of the Thesaurus approach

To assess the overall performance of our more elaborate query, we generate an uncertainty ratio of how many words of any given article are contained in the query as a fraction of total words in the article. Based on this ratio, we can then define a threshold x above which an article is classified as talking about environmental policy uncertainty (e.g. EnvPU = 1 if uncertainty ratio ≥ x%). Next, we compute precision and recall of all classification rules, i.e. for

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

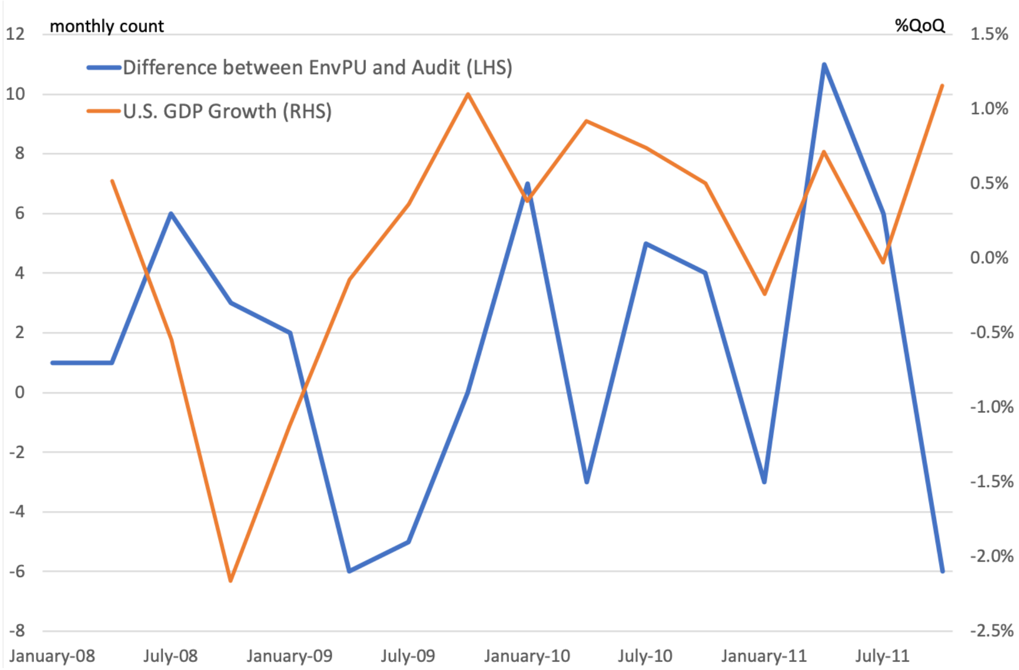
Table D.2: Keywords for predicting our EnvPU index

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

uncertainty thresholds between 0.1% and 3%. Intuitively, precision is highest the higher (more restrictive) the uncertainty threshold is but this comes at the cost of a lower recall. We choose an uncertainty threshold of 0.4% for our baseline prediction rule with a precision of 0.55, recall of 0.53 and an F1 measure of 0.54. These numbers represent a significant improvement compared to the performance of the 'uncertain*' query. However, they are still inferior as compared to our SVM based algorithm.

Extra material on the audit

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



EnvPU and elections

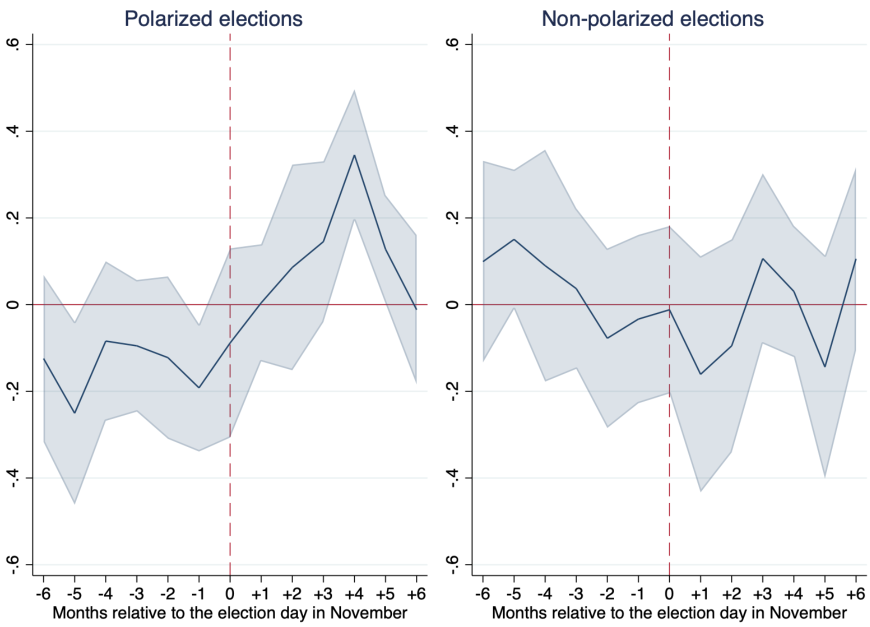
Table D.3: US Presidential Elections, 1984 onwards

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

*: Based on data in two states.

**: Assumed to be polarized.

Figure D.2: EnvPU and presidential elections



This figure presents the coefficients on dummies for six months before and after a presidential election depending on whether the election is polarized (left panel) or not (right panel) (i.e., Columns (2) and (3) in Table D.4). The value of these coefficients reflects the level of EnvPU during each of these months relative to the rest of the sample. The shaded area is the 90% confidence interval.

Table D.4: EnvPU and Elections

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

The table presents results of an OLS regression using Equation (4.1). The sample period is January 1983 to December 2018. The dependent variable is the logarithm of the EnvPU index. Column (2) restricts the sample to polarized elections and Column (3) to non-polarized elections. Robust standard errors are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Additional statistics and results: Venture Capital

Descriptive statistics - firm-level estimations and VAR

Table D.5 reports descriptive summary statistics for the main continuous variables used in our study of VC investments for the period between January 1998 and March 2019. The first panel shows the variables that are common to both the VAR and Panel analysis, taken from our monthly VAR dataset. The second panel displays statistics about the monthly number of VC deals in the United States that we use in our VAR analysis. Finally, the last panel reports the VC-related variables that we use in our panel analysis. The discrepancy between the number of observations reported in the last panel and in Table 3.3 comes from the fact that, in our analysis, we drop all the observations where the age variable is negative as it most likely indicates an error in the data. In other words we drop around 30,000 observations that were recorded before the startups' official founding date.

Table D.5: Summary statistics for venture capital analysis

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

Additional results and statistics on stocks

Table D.6: Summary statistics for stock return analysis

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

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

Matthias van den Heuvel

526 E 20th Street, NY 10009 | matthiasvandenheuvel@gmail.com | +1 (649) 510 4119 | Swiss and British

Economist specialized in trade, environmental economics and finance with strong analytical and communication skills. Through my work in academia, private banking and international organizations, I have experience working in diverse settings. I will complete my PhD in Economics with a focus on the financing of clean technologies during the first quarter of 2022. I am fluent in English and French.

Experience

Swiss Federal Institute of Technology (EPFL) — Lausanne, Switzerland
PhD Researcher, Chair of Innovation and IP Policy — Feb 2018 – March 2022

- Conduct research on the financing of investments in clean technologies funded by the Swiss National Research Programme "Sustainable Economy: resource-friendly, future-oriented, innovative" (NRP 73)
- Teach and develop material for Master courses: Introduction to Econometrics (on R) and Economics of Innovation and IP (recipient of the 2021 EPFL's MTE Best Teaching Assistant Award).
- Present my research and its policy applications to academic audiences and project partners (e.g. OECD, Venture Kick)

Edmond de Rothschild Private Bank — Geneva, Switzerland
Economist, Economic Research Team — Jan 2016 – Jan 2018

- Drafted articles on various economic issues, including the housing market in the Eurozone, the European Central Bank and Swiss National Bank's monetary policies or the impact of slow wage growth on the Japanese economy
- Built macroeconomic forecasting models for Switzerland, Japan and the Eurozone
- Contributed to a bi-annual macroeconomic forecasting document to inform the bank's board, advisors and clients
- Gave presentations and recommendations to clients and bankers on issues related to the Swiss economy
- Supported the Chief Economist Mathilde Lemoine in the production of media articles, interviews and presentations

World Trade Organization — Geneva, Switzerland
Researcher, Economic Research and Statistics Division — Mar – May 2015

- Drafted a section of the World Trade Report on Trade Facilitation
- Participated in the elaboration of a trade course for economists in Bangkok

Center for International Environmental Studies — Geneva, Switzerland
Research Assistant — Jun – Aug 2014

- Produced a report on the international coal market
- Estimated the impact of an EU carbon border tax on South Africa's coal industry

Education

Swiss Federal Institute of Technology (EPFL) — Lausanne, Switzerland
PhD in Economics. Dissertation: *Financing investments in clean technologies* — Feb 2018 – Feb 2022

- Attended the Swiss Program for Beginning Doctoral Students in Economics (Gerzensee, 2018-2019)
- Achieved an overall GPA of 5.6/6 in Microeconomics, Macroeconomics and Econometrics

Yale University Law School & Department of Economics — New Haven, United States
Exchange Semester with a Scholarship from the Bakala Foundation — 2014 – 2015

Graduate Institute of International and Development Studies — Geneva, Switzerland
Master in International Economics: Valedictorian, GPA 5.64/6 — 2013 – 2015
Master's Thesis: *Collateral Damages or Benefits; The impact of UN sanctions on the Neighboring Countries of the Target*

University of New South Wales — Sydney, Australia
Exchange Semester with a Scholarship from the Boninchi Foundation — 2013

University of Geneva — Geneva, Switzerland
Bachelor in International Relations: top 1% among economics majors, GPA in economics 5.8/6 — 2010 – 2013

Publications

van den Heuvel, M., et al. "COVID-19: Insights from innovation economists." Science and Public Policy 47.5 (2020): 733-745

de Rassenfosse, G., van den Heuvel, M. (2021). Certification or Cash Prize: The Heterogeneous Effect of Venture Competitions. Available at SSRN, under review.

Noailly, J., Nowzohour, L., van den Heuvel, M. (Forthcoming). Heard the News? Environmental Policy and Clean Investments

Noailly, J., Nowzohour, L., van den Heuvel, M. (Forthcoming). Environmental Policy Uncertainty: Measurement and Impact on Clean Markets

Popp, D., van den Heuvel, M. (Forthcoming). The role of Venture Capital and Public Investors in Financing Clean Energy

Selected Contributed Talks

Presented *Certification or Cash Prize: The Heterogeneous Effect of Venture Competitions*, at the SEI2019 Consortium and the 2021 EPFL Virtual Innovation Seminar.

Presented *Heard the News? Environmental Policy and Clean Investments*, at the 2020 NBER Energy Conference, the 2021 EAERE Conference and the 2021 European Economic Association Conference.

Qualifications and Interests

Computer skills: Stata, Python, R, Eviews, LaTeX
Technical skills: Data Analysis, Econometrics and Machine Learning
GRE: 170Q, 169V and 5.0 AWA
Driving Licenses: A, B, C, D1, BE, CE, D1E (Driver in the Swiss Army) and Open Water Diving (SSI)
Interests: Scuba diving, playing Basketball and Magic the Gathering, and reading Books.

Languages

French C2; English C2; German B2 (in the process of learning German); Italian A2