Green Entrepreneurship

Wan-Chien Chiu[a1], Shenje Hshieh[b2], Po-Hsuan Hsu[c1], and Geoffery Zheng[d3]

[1]National Tsing Hua University
[2]City University of Hong Kong
[3]NYU Shanghai

March 2026

Abstract

By constructing a comprehensive dataset of startups and their patenting and trademarking activities, we track the evolution and landscape of green entrepreneurship in the U.S. over the past four decades. We observe strong growth in the creation of green intellectual properties (IPs) by startups in aggregate: the proportion of startups with any green IPs increases from 2% in the 1980s to 7% in the 2010s. Specifically, the fraction of green patents (trademarks) filed by startups grew from 5% (2%) in the 1980s to 23% (5%) in the 2010s. We also observe significant cross-state variation in the growth of the proportion and count of startups with green IPs and startups having self-curated business profiles with high frequencies of environment-related keywords. We further examine how the surge of green entrepreneurship could be explained by education, policy incentives, and public attention to climate issues. We document significant increases in the number of green startups and the average startup's proportion of green IPs when (i) a green research institute is established in a local university, (ii) local community receives more green investment from the American Recovery and Reinvestment Act, and (iii) natural disasters impact neighboring counties.

[a]wan-chien.chiu@mx.nthu.edu.tw
[b]shshieh@cityu.edu.hk
[c]pohsuanhsu@mx.nthu.edu.tw
[d]geoff.zheng@nyu.edu. We thank Michael Ewens, Michelle Lowry, and Matt Marx for helpful discussion. We thank Shihe Li for his great research assistance. All errors are our own.

1 Introduction

Entrepreneurs play an important role in the economy not only through creative destruction but also because they are motivated by factors beyond private pecuniary gains. Entrepreneurs often incorporate the pursuit of self-fulfillment, the joy of creating things, and independence into their venture objectives.[1] Among all subjective incentives, social preferences may be one force that drives entrepreneurs to change the society.[2] In this paper, we first present the landscape of an expanding group of "green startups" that focus on business opportunities related to environmental issues in the past few decades. We then evaluate their development of green technologies and products, and further examine the motivating forces behind their emergence in both economic contribution and technological development.

Economic growth and environmental protection were first considered conflicting targets in the 1970s.[3] Later, those two seemingly opposing goals were reconciled by a new direction of "Sustainable Development" that meets current people's needs without compromising the ability of future generations to meet their own needs, a concept was developed and promoted by the United Nations' "Our Common Future" report in 1987.[4] This consensus was shared by several pioneering entrepreneurs who founded Patagonia (1973), the Body Shop (1976), and Ben & Jerry's (1978), which led a long trend of social enterprises that include green startups.

To present the evolution and landscape of green startups, we first collect U.S. startups' financing and accounting data from Pitchbook and their founders' and employees' profiles from Revelio Labs. We then match those startups to the assignees of patents and trademarks from the U.S. Patent and

1. In his discussion of entrepreneurial motives (Schumpeter 1934, p. 68), he concludes *"Nevertheless it is true that the second and third groups of entrepreneurial motives may in principle be taken care of by other social arrangements not involving private gain from economic innovation. What other stimuli could be provided, and how they could be made to work as well as the "capitalistic" ones do, are questions which are beyond our theme. They are taken too lightly by social reformers, and are altogether ignored by fiscal radicalism."* Other non-pecuniary incentives include time flexibility (Hurst and Pugsley 2011), preferences for non-routine tasks (Levine and Rubinstein 2017), and preference for risk (Sauermann 2017; Bernstein, Townsend, and Xu 2023). Prior studies also examine how such non-pecuniary utility is revealed (Bellon et al. 2021).

2. An extensive list of experiment, survey, and empirical studies have collectively supported individuals' social preferences (Preston 1989; Sen and Bhattacharya 2001; Bhattacharya and Sen 2003; Tonin and Vlassopoulos 2010; Bénabou and Tirole 2009).

3. For instance, the Conference on the Human Environment hosted by the United Nations in 1972 was regarded as the first global initiative that highlighted the pressing need of environmental protection and set up plans for environmental coordination globally. (Source: https://thesustainablemag.com/environment/the-history-of-sustainable-development-goals-sdgs) In the same year, the influential report "The Limits to Growth" by the Club of Rome and the Massachusetts Institute of Technology (MIT) presented model predictions of resource depletion and potential collapse of civilization given the growth trajectories of population, industrialization, resource depletion, and pollution back then.

4. https://www.brundtland.co.za/wp-content/uploads/2022/08/Brundtland-Report-1987-Our-Common-Future.pdf

Trademark Office (USPTO), which allow us to track those startups' development of environment-oriented technologies and product lines. In particular, we follow the guide of European Union Intellectual Property Office (EUIPO) to classify green trademarks and employ three ways to identify green patents following Bolton, Kacperczyk, and Wiedemann (2022), WIPO, and Sun, Tian, and Yang (2024) – all of these enable us to track how startups perform in terms of green patents and trademarks.

Our data endeavor result in 1,052,347 patents (195,552 are green patents) and 395,932 trademarks (14,605 are green marks) created by 88,321 startups in our sample. At the aggregate level, we illustrate a steady increasing trend in green innovations. In terms of patents, we observe a stable increase in the percentage of green patents filed from 15% in 1980 and to 25% in 2023 among all startups. Such a pattern is also observed in the percentage of green trademarks filed, rising from 2% in 1980 to 5% in 2023.

In order to capture green entrepreneurship broadly, we define a "green startup" as either (i) a startup that has created green intellectual properties or "green innovations" as measured by green patent and trademark filings within the first five years of founding or (ii) a startup whose self-curated business profile in Revelio Labs contains a high proportion of environment-related keywords.

We designate a startup from the first group as a "green-innovation startup". Our sample contains 9,070 green-innovation startups. In the 1980s (the earliest period of our sample), we identify only 176 of such startups (1.6% of all startups). This number increases to 4,090 for those founded in the 2010s towards the end of our sample, accounting for 4.7% of all startups. In terms of geographic distribution, more green-innovation startups are formed in states with larger populations (e.g., California, New York, Texas, Massachusetts, and Pennsylvania). More interestingly, the proportion of such startups also increases in these states across time, suggesting that the rise of green entrepreneurship is a general pattern across all states. In terms of sectorial distribution, the energy equipment sector naturally has the highest percentage of green-innovation startups of 30.8%. The semiconductor sector is ranked the second with 21.7%. Green-innovation startups, on the other hand, make up only 5.5% of all startups in the software sector, which is the largest sector in our sample.

We designate a startup from the second group as a "self-labeled green startup" if the fraction of environment-related keywords in its business description is in the top quartile of the entire sample of startups. The Bag-of-Words for environment-related keywords is the union of those from Baier, Berninger, and Kiesel (2020), Lin et al. (2024), and Pencle and Mălăescu (2016). We confirm that

this relative definition strongly correlates with green patenting and trademarking. We find that these self-labeled green startups file their first green patents earlier and are also more likely to continue to do so through their life cycle compared to other startups. For example, self-labeled green startups are twice as likely to file their first green patent within 3 years of inception. In addition, self-labeled green startups on average file 0.04 green patent within 4 years of inception, which is twice as much as the other startups. A similar yet even greater contrast is also observed in green trademarking: self-labeled green startups are four times more likely to file their first green patent within 3 years of inception. Moreover, within 1 year of founding, a self-labeled green startup on average files 0.13 green trademarks compared to 0.03 for the other startups. Overall, our analyses of the likelihood and frequency of green patenting and trademarking between self-labeled green startups and others confirm that the former group indeed "walk their talk".

In the third part of this paper, we examine potential factors that may explain green startups' emergence; in particular, we focus on education, policy incentives, and public attention to climate issues.

From the perspective of education, we first construct a dataset of the establishments of green institutes and centers at U.S. universities. As universities serve as an incubator for basic and applied sciences as well as an information hub for researchers and industry partners, we expect the foundation of environment research institutes to explain the growth in local green innovations and startups. In a staggered difference-in-differences setting, we find a significant increase in the proportion of green trademarks among new startups after the founding of such institutes in a geographic area. In terms of economic significance, when startups are founded in commuting zones with green research institutes, their green trademark ratio increases by 20.5% relative to the unconditional mean. In addition, the number of green-innovation startups in commuting zones with green research institutes grows by 25% than those without. Our results offer strong evidence supporting an intuitive relation between environment- or climate-related education and green entrepreneurship.

From the perspective of policy initiatives, we consider one of the largest government stimulus packages called the American Recovery and Reinvestment Act (ARRA) that allocated over $350 billion of direct government spending and an additional $260 billion in tax reductions (Aldy 2013). In particular, $51 billion was allocated to the Environmental Protection Agency (EPA) and Department of Energy (DoE) for green investment in 2009, the first year.[5] To measure the heterogeneity

5. These investments primarily target clean energy initiatives (e.g., renewable energy generation, smart grid technologies, and advanced battery storage) and environmental protection programs (e.g., EPA Superfund remediation of contaminated sites, wastewater treatment infrastructure, and diesel emissions reduction efforts). Popp et al. (2021)

in spatial allocation of green ARRA funding, we calculate the per-capita dollars of ARRA-related green investment in each commuting zone. We find that new startups in areas receiving more green investment focus more on green trademarking after 2009; when startups are founded in commuting zones where ARRA funding per person increases from the 25th to the 75th percentile, the average proportion of green trademarks filed increases by 37.9%, suggesting the launch of relatively more environment-related product lines. We also find that the inter-quartile range change in ARRA funding had an effect on startup founding frequency, adding an average of one additional green startup per commuting-zone-year in the post-ARRA period. These results highlight the important role of government policy initiatives in green entrepreneurship.

Finally, from the perspective of public attention, we use natural disasters (e.g., floods, hurricanes, tornadoes, wildfires, etc.) in adjacent counties to measure increased attention to climate issues (Chiu et al. 2024). We consider a county being treated in a year if one of its neighboring counties was impacted by natural disasters (but itself was not). Startups in treated areas file relatively more green trademarks, as the fraction of green trademarks increases by 0.2 to 0.6 percentage points in the years following the natural disaster. More importantly, the number of new green-innovation startups in treated areas grows by 29.2% in the year following the natural disaster.

Our analyses of potential contributing factors to green entrepreneurship and the associated green innovations suggest that the increase of green startups follows the establishment of green research institutes in U.S. universities, the injection of green investment from government policies, and the observation of natural disasters in neighboring geographic areas. While these factors also influence green trademarking by new startups, we do find any effect on green patenting. This pattern is intuitive given that startups' top priority is on commercialization. This is consistent with green research institutes developed in local universities to not only enhance technological innovation but also commercialization for local entrepreneurs. Similarly, public green investment and attention to climate issues call for immediate products and/or applicable and commercializable inventions, which are more likely reflected in green trademarks than in green patents.

This study contributes to the literature on entrepreneurship, innovation, sustainable growth, and environmental policies. First, we add to the entrepreneurship literature by highlighting an important group of entrepreneurs that have social preferences or is seeking environment-related business opportunities. Through the analysis of self-curated business profiles, we show that those entrepreneurs indeed engage more in developing green technologies and launching green products,

provide a detailed discussion.

validating our definition of green startups and their corresponding actions. Our effort in linking startups to patents and trademarks and classifying green innovations also solves the problem of how to measure the ESG activity of startups, a long-standing challenge in social enterprises.[6] We also highlight the role of education, government-led investment, and public attention in local entrepreneurship from the perspective of sustainability.[7] Therefore, this study advances the understanding of startup performance metrics and the underlying drivers of entrepreneurial activity.

Second, we construct new datasets and novel measures for the literature on sustainable growth. Different from prior studies that focus on big firms — especially heavy emitters (Akey and Appel 2020; Xu and Kim 2021; Duchin, Gao, and Xu 2024; Hsu, Li, and Pan 2025),[8] we investigate the operations and investment decisions of startups. Green startups and startups with green innovations constitute another significant market force in sustainability- or climate-related businesses. Despite their relatively small size at present, their potential in both technology and product frontiers has long-term implications and thus calls for more research. It is also interesting to note that startups' green trademarks react stronger than their green patents, which highlights their more prominent footage in the product space instead of the technology space. Overall, our comprehensive dataset that covers various aspects of green entrepreneurship creates a new avenue of sustainability research.

Third, we contribute to the innovation literature by presenting the evolution and landscape of green innovations created by startups. Unlike prior studies on green innovation by big firms (such as publicly-listed ones) and specific industries (Brown, Martinsson, and Thomann 2022; Sautner et al. 2023a, 2023b; Cohen, Gurun, and Nguyen 2026),[9] we collect startups' environment-related patents and trademarks and document strong growth among those young firms over the past few decades. Also, by linking publicly available trademark records to startups, we illustrate a new method to identify new ventures' activity in the product space,[10] which enables future researchers to examine how entrepreneurs convert new technologies into product lines.

6. As rating agencies such as MSCI and Refinitiv usually do not rate small and medium-sized firms, it is challenging for us to use existing ESG dataset to measure startups' green innovation. Thus, prior studies rely on surveys (e.g., Horbach (2008)).

7. Prior studies have examined how entrepreneurship is influenced by financial conditions and risk considerations (Hurst and Lusardi 2004; Hombert et al. 2020; Bellon et al. 2021), as well as culture and social factors (Guzman, Oh, and Sen 2020; Engelberg et al., n.d.)

8. An exception is Bellon et al. (2021) that uses all establishments from the Longitudinal Business Dataset (LBD). Moreover, prior studies have also examined how shareholders, bondholders, and banks influence corporate missions (e.g., Krüger (2015); Starks, Venkat, and Zhu (2026); Krueger, Sautner, and Starks (2020); Flammer (2021); Kacperczyk and Peydro (2021); Hsu, Li, and Tsou (2023); Ivanov, Kruttli, and Watugala (2023)).

9. In addition, prior studies also attempt to develop different measures for firms' green revenues (Klausmann, Krueger, and Matos 2024), green products (Meier et al. 2022; Chiu et al. 2024), and green human capital (Darendeli, Law, and Shen 2022; Sautner et al. 2023a).

10. Ewens and Marx (2023) present how to use patenting records to measure entrepreneurial activity.

Finally, this study offers new insights on government policies related to environmental issues and the development of small and medium-sized enterprises (SMEs). Prior research on environmental and climate policies has covered pollution penalty, mandatory disclosure, carbon tax, green loans, and tax credits for green investment.[11] However, the discussions on governments' direct investment and initiatives for green entrepreneurship are lacking in the past. Our results on the impact of ARRA green funding point to a spillover from governments' green investment to local startups, suggesting possible complementarity of environmental policies.

While prior studies on SME policies extensively examine policy tools that promote technological innovation and economic growth,[12] they pay limited attention to green innovation and environmental contributions. From the perspective of sustainable growth, the creative destruction induced by green startups is particularly important, as incumbent firms often replace existing products and operations only slowly. Our evidence highlights the importance of SME policies that recognize the environmental contributions of these firms in promoting sustainable growth.

2 Data Sources and Sample Construction

We first discuss how we merge various datasets to create a unified dataset of startups' intellectual property portfolios. We then discuss our data sources and variables for startup characteristics, trademark applications and grants, and patent applications and grants. We then discuss how we classify trademarks and patents as green and non-green. Finally, we discuss the data for the three factors (green research institutes, government green investment, and natural disasters) we use to explain green innovations and entrepreneurship.

2.1 Linking across Datasets

Linking Pitchbook and Revelio We join Pitchbook with Revelio on the basis of name and online presence. For each company in the two datasets, we clean the reported website and LinkedIn profile URLs. We keep firms which are uniquely linked by these two URLs, as well as firms which

11. An incomplete list of prior studies includes Newell, Jaffe, and Stavins (1999), Popp (2002, 2003), Popp, Newell, and Jaffe (2010), Acemoglu et al. (2012), Acemoglu et al. (2016), Aghion et al. (2016), Chen, Hung, and Wang (2018), Downar et al. (2021), Greenstone et al. (2022), and Tomar (2023).

12. The effectiveness of SME policies have been extensively discussed in the past, including government-sponsored venture capitalists (VCs) (Lerner 2002, 2009; Brander, Du, and Hellmann 2014), R&D subsidy (Gans and Stern 2003; Howell 2017), government procurement (Feldman and Kelley 2003), public financing programs (Toole and Turvey 2007), accelerators (Gonzalez-Uribe and Leatherbee 2017), government-sponsored R&D alliances (Doblinger, Surana, and Anadon 2019), and public R&D investment (Azoulay, Graff Zivin, and Manso 2011; Babina et al. 2023; Azoulay, Heggeness, and Kao 2025). Howell (2024) offers a contemporary review of related policies.

are uniquely linked by one of the two URLs and missing for the other. From Pitchbook, we have a total of 278,037 startups (VC and PE firms). We managed to match 197,670 of them to 197,497 firms in Revelio.

Linking Trademarks to Revelio We link USPTO trademark data to the Revelio dataset by mapping corporate trademark owners to their unique firm identifiers (RCID). The process begins by filtering the USPTO dataset to include only corporate entities, excluding individual owners. We then harmonize the entity names across both datasets to ensure structural consistency. A fuzzy matching algorithm is subsequently employed to identify potential matches, which are then subjected to a rigorous manual validation process using Google search and official company websites to ensure accuracy. We identify 395,932 trademarks filed between 1973 and 2023 by 79,139 unique Pitchbook startups (79,003 firms from Revelio).

Linking Patents to Revelio We make use of a patent-to-firm link table provided by Revelio to assign patents to Revelio firm identifiers (RCID).[13] We identify 1,052,347 unique patents filed between 1955 and 2023 by 28,657 unique startups (28,608 firms from Revelio).

2.2 Startup Characteristics

We use data on startups from Pitchbook and Revelio Labs. Pitchbook contains data on startups' formation, industry, online presence, and trade names. From Revelio Labs, we get individual profile as well as company profile data. The data is collected from LinkedIn. From individual profiles we obtain their self-reported employment history and education history. From company profiles we obtain their online presence as well as a voluntary description of their business. Our sample comprises roughly 175,000 firms, the majority of which are founded between 1998 and 2021. For a large subset of these firms, we have data on industry, location, and a business description.

We begin by presenting a time series of startup formation in Figure 1. Our sample begins in 1980 and ends in 2023. We observe a dramatic increase in the rate of startup formation over the sample, with some attrition towards the end of the sample.[14] Not only is the startup creation rate increasing, we also observe an increase in early-stage innovation.

Table 1 reports summary statistics of the startups that comprise our sample of startups (restricted to those with non-missing found year) merged with business description, patent, and trade-

13. We note that the Revelio Labs subscription on WRDS does not contain this linking table.
14. This truncation bias is common in datasets of young firms.

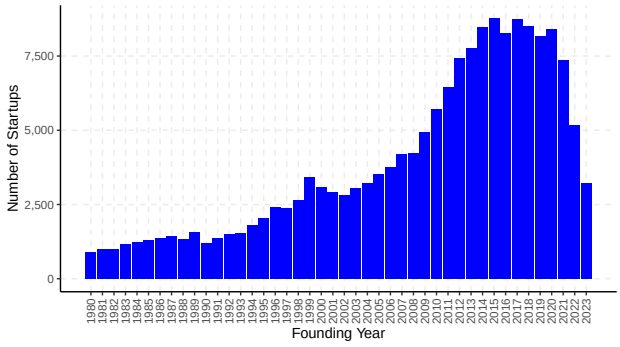


Figure 1: The number of startups founded each year. Source: Pitchbook

mark data. The average startup has 6 granted patents and 2 registered trademarks. In terms of green innovations, the average startup has 1 green patent and 0.08 green trademarks. About 9.3% of our sample of startups have at least 1 green patent or trademark since founding. That number falls to 5.2% if we restrict to green patents and trademarks filed within 5 years of founding. In our sample, 90% of startups provide business descriptions, which serve as another data source for quantifying the "greenness" of their business models.

Table 1: Descriptive Statistics of Startups

	N	Mean	SD	p5	p25	p50	p75	p95
Year Founded	174,926	2003	22.977	1959	1998	2011	2017	2021
Patents Granted	174,926	5.800	251.577	0.000	0.000	0.000	0.000	7.000
Green Patents Granted	174,926	1.057	38.205	0.000	0.000	0.000	0.000	1.000
Trademarks Registered	174,926	2.280	15.382	0.000	0.000	0.000	2.000	8.000
Green Trademarks Registered	174,926	0.082	0.725	0.000	0.000	0.000	0.000	0.000
Have Green Innovation	174,926	0.093	0.290	0.000	0.000	0.000	0.000	1.000
Have Green Innovation within 5 Yrs	174,926	0.052	0.222	0.000	0.000	0.000	0.000	1.000
Green Keyword Ratio	158,374	0.055	0.047	0.000	0.025	0.047	0.074	0.141
Has Business Description	174,926	0.905	0.293	0.000	1.000	1.000	1.000	1.000
Has Industry Classification	174,926	1.000	0.013	1.000	1.000	1.000	1.000	1.000

Notes: Year Founded is the founding year of the startup in the Pitchbook database. Patents Granted is the number of patents issued to a startup throughout its observable lifetime. Green Patents Granted is the number of green patents issued to a startup during its observable lifetime. Trademarks Registered is the number of trademarks registered a startup during its observable lifetime. Green Trademarks Registered is the number of green trademarks registered by a startup throughout its observable lifetime. Have Green Innovation is an indicator for whether a startup has filed a patent (eventually granted) or trademark during its observable lifetime. Have Green Innovation within 5 Yrs restricts the previous variable to patents and trademarks filed within 5 years of startup founding. Green Keyword Ratio is the fraction of green keywords in a startup's business description text. Has Business Description is an indicator for whether the startup has a business description. Has Industry Classification is an indicator for whether the startup has industry classification information.

In Table 2, we present the industry distribution of startups in our sample in all sample years and in each decade (1980s, 1990s, 2000s, and 2010s). Software has the largest representation in our



Figure 2: Word Cloud of Environmental Terms

Notes: Environmental terms are taken from Baier, Berninger, and Kiesel (2020), Lin et al. (2024), and Pencle and Mălăescu (2016). Size of words is determined by their frequency in the LinkedIn corpus.

sample with 43,383 startups across all years. Commercial services industry is ranked the second with 28,074 startups, and commercial products industry is ranked the third with 11,692 startups. In the same table we break down the statistics by decades, starting in the 1980's and ending in the 2010's. The overall number of startups founded each decade grow by 778% from the 1980's to the 2010's. Among well-populated industries, defined as industries for which we consistently observe over 50 firms in each decade, Software (12,082 startups founded in the 2010's), Healthcare Technology (6,119 startups founded in the 2010's), and Media (2,407 startups founded in the 2010's) are the three fastest growing industries. The growth rates in these three industries are 1,297%, 1,269%, and 1215.3%, respectively, and are almost double the overall sample growth rate. Industries which exhibit relatively slow growth are Commercial Transportation (575%), Transportation (594%), Capital Markets (603%), and Insurance (519%). Although software is both the largest and the fastest growing industry in our sample, we observe no overall correlation between industry size and growth rates.

Business Descriptions and Environmental Terms We use a Bag-of-Words approach to classifying startups as green. The dictionary for environmental terms is the union of three dictionaries taken from Baier, Berninger, and Kiesel (2020), Lin et al. (2024), and Pencle and Mălăescu (2016). For each startup, we take the business description in Revelio and compute the fraction of words that are labeled as environmental. In Figure 2 we construct a word cloud of the words in the environmental keywords dictionary that appear most frequently in the Revelio corpus.

Table 2: Sectoral Distribution of All Startups by Founding Decade.

Industry Group	All Years	1980s	1990s	2000s	2010s
Agriculture	809	46	62	104	366
Apparel and Accessories	2,449	117	175	378	1,034
Capital Markets/Institutions	1,645	173	277	412	1,043
Chemicals and Gases	1,131	128	139	207	798
Commercial Banks	720	63	80	217	651
Commercial Products	11,692	1,630	1,504	1,798	9,324
Commercial Services	28,074	2,721	4,214	6,185	17,739
Commercial Transportation	1,274	171	177	219	984
Communications and Networking	2,406	186	665	737	1,848
Computer Hardware	3,272	213	297	634	1,460
Construction (Non-Wood)	183	8	26	29	137
Consumer Durables	4,325	269	373	636	2,010
Consumer Non-Durables	8,069	377	533	1,020	3,021
Containers and Packaging	784	92	113	108	641
Energy Equipment	1,108	53	71	364	608
Energy Services	1,379	69	124	363	724
Exploration, Production and Refining	941	66	71	314	559
Forestry	172	25	20	22	141
Healthcare Devices and Supplies	5,791	249	565	1,378	2,650
Healthcare Services	9,892	890	1,576	2,314	6,119
Healthcare Technology Systems	4,515	96	286	727	1,314
IT Services	4,055	301	877	1,307	2,746
Insurance	2,872	421	394	474	2,185
Media	5,638	183	437	1,110	2,407
Metals, Minerals and Mining	346	29	42	62	249
Other Business Products and Services	903	104	113	193	659
Other Consumer Products and Services	15	1	1	0	3
Other Energy	127	5	9	32	71
Other Financial Services	3,805	214	424	671	1,661
Other Healthcare	20	1	2	1	4
Other Information Technology	15	0	0	3	5
Other Materials	261	11	25	47	116
Pharmaceuticals and Biotechnology	7,454	241	642	1,702	2,942
Restaurants, Hotels and Leisure	3,217	232	357	575	1,725
Retail	2,163	157	246	372	1,292
Semiconductors	788	61	128	257	509
Services (Non-Financial)	6,676	417	624	1,108	3,145
Software	43,383	865	2,651	6,687	12,082
Textiles	161	13	22	20	115
Transportation	2,049	210	219	316	1,248
Utilities	319	11	33	58	217
Total	174,898	11,119	18,594	33,161	86,552

2.3 Classifying Green Innovations

Classifying Green Trademarks We employ a novel text-based classification methodology to identify green trademarks registered with the USPTO between 1980 and 2023. This approach is based on European Union Intellectual Property Office (2021), which originated from a 1996 initiative where environmental experts identified terms linked to environmental improvement (Annex 1). To ensure longitudinal consistency across historical and future filings, the EUIPO refined this inventory into 375 standardized green rules using a hybrid approach of machine learning and expert intervention (Annex 2).[15] Following the precedent set by Chiu et al. (2024) for U.S. publicly traded firms, we operationalize this taxonomy by extracting and parsing the Goods & Services (G&S) descriptions from USPTO records.[16] Around 3.69% of all trademarks included in our sample are classified as "green". For a more detailed description of the classification process, refer to Online Appendix A.1.

Classifying Green Patents We employ three classification methodologies to identify green patents. The first methodology is from Bolton, Kacperczyk, and Wiedemann (2022) and is based on the CPC code. We refer to these as the "Green: BKW" patents. The second methodology is based on the IPC Green Inventory classification from the WIPO.[17] We refer to these as the "Green: WIPO" patents. The third methodology is from Sun, Tian, and Yang (2024) and follows the 'Y02/Y04S' tagging scheme proposed by the European Patent Office in 2012. We refer to these as the "Green: Robust" patents. Finally, we construct the union of these classifiers, so that any patent classified as green under one of the three methodologies is included. We refer to these as the "Green" patents. Around 18.58% of all patents included in our sample are classified as "green". For a more detailed description of the classification process, refer to Online Appendix A.2.

15. For example, Ref. 224's green expression is +solar +energy. This rule indicates that, for a USPTO trademark to be identified as green, its goods-and-services description must contain both "solar" and "energy" anywhere in the description. As another example, Ref. 73's green expression is +electric +vehicle −cigarette −door −horn −lock −sunroof −alternator −alarm −temperature −theft −antitheft −washers −7 −37. This rule indicates that, for a USPTO trademark to be identified as green, its goods-and-services description must contain both "electric" and "vehicle" anywhere in the description, but it must not contain any of the excluded terms (e.g., "cigarette," "door," "horn," "lock," "sunroof," "alternator," "alarm," "temperature," "theft," "antitheft," or "washers"). In addition, the rule excludes specific NICE Classes (in this rule, Class 7 or 37) to prevent false positives in non-green industrial contexts.

16. A uniform preprocessing pipeline is applied to both the trademark Goods & Services descriptions and the 375 green expressions defined in the EUIPO Green Trademark Taxonomy. The procedure involves removing punctuation and filtering out low-semantic function words (stop words) such as "a" or "is" using the Python packages `stop-word` and `pysnowball`. Following tokenization, hyphenated compounds are standardized using a dual-rule approach to minimize format variance: hyphens are both removed (e.g., "self-balancing" to "selfbalancing") and replaced with spaces (e.g., "electric-motor" to "electric motor"), with both transformations applied to every compound term.

17. We first map patents CPC codes to IPC codes.

2.4 Factors for Green Entrepreneurship

Green Research Institutes We construct a dataset of green institute and center openings at U.S. universities ("green research institutes"). To do this, we begin with a list of 1,726 US higher-education institutions, made available at unirank.org. From this list, we select the top 200 universities by rank and use the GPT 5.2 model, via the OpenAI API, to identify environmental, climate, and sustainability-focused academic units at these schools along with the year of founding.[18] Our sample consists of 272 green research institutes founded between 1980 and 2023 and located in 82 unique commuting zones. For additional details, refer to Online Appendix A.3.

Green ARRA Spending We use data from the NBER American Recovery and Reinvestment Act (ARRA) dataset on Federal Spending by commuting zone. We use geographic variation in ARRA funding as a quasi-random shock to green funding availability and analyze its effects on green startup formation and green innovation. To isolate the component of ARRA funding that is related to green initiatives, we follow the methodology of Popp et al. (2021). Green spending is spending that is intended for immediate job creation and allocated to: (i) the Department of Energy (DoE) for renewable energy and energy efficiency retrofits; and (ii) the Environmental Protection Agency (EPA) for brownfield redevelopment and renewable energy. We summarize aggregate ARRA spending amounts in Online Appendix Table B.2.[19]

Climate Disasters We use the SHELDUS (Spatial Hazard Events and Losses Database for the United States), a database maintained by Arizona State University to identify natural disasters (e.g., floods, hurricanes, tornadoes, wildfires, etc.) and their economic impact at the county level across the United States.[20] Specifically, the dataset includes detailed information on property damage, crop damage, injuries, and fatalities. The data is compiled from sources like the National Weather Service, the U.S. Geological Survey, and other federal agencies, covering events from 1960 onward. In this paper, we use SHELDUS to capture climate-related natural disasters as a proxy for heightened managerial awareness of environmental risks.

To isolate the awareness effect from the direct economic costs of natural disasters, we construct a county-level indicator variable equal to 1 if the county is adjacent to at least 1 disaster-affected

18. Unirank ranks accredited institutions offering bachelor's and/or postgraduate degrees through primarily face-to-face learning in on-site facilities. The ranking is based on four independent web metrics.

19. The ARRA data are obtained from National Bureau of Economic Research, "American Recovery and Reinvestment Act (ARRA) Federal Spending Detail 2009–2012," available at https://data.nber.org/data/ARRA/.

20. See https://sheldus.asu.edu/.

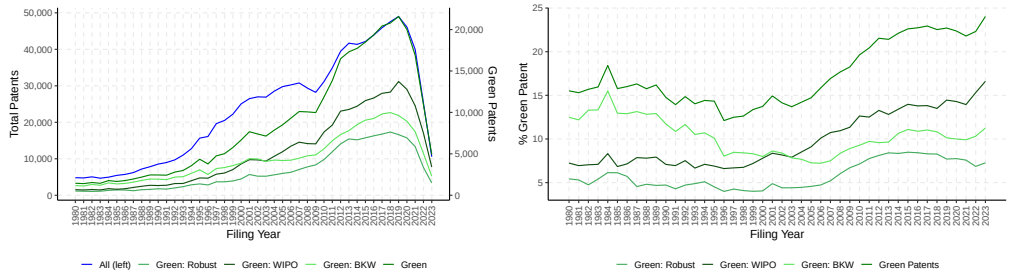
county with total annual property and crop damage exceeding 1 million USD (adjusted to 2023 dollars).

3 Green Innovations by All Startups

In this section, we analyze the green innovative activity of all startups. In Figure C.1 in Online Appendix we present a time-series of the number and fraction of startups that engage in intellectual property activity. The blue line is the fraction of startups founded each year which go on to file a patent or trademark at some point in their lifecycle. This line has remained around 50 percent throughout the sample, with some attenuation at the end owing to the delayed release of intellectual property filings. This line suggests that intellectual property has always been an important part of startup innovation. However, we see that the fraction of startups engaging in *early stage* (within 5 years) innovation is increasing.

Motivated by the marked increase in early-stage filings, we analyze the green filings of startups. In Figure 3a, we present the total number of patents (on the left vertical axis) and green patents (on the right vertical axis) filed against the filing year on the horizontal axis. We observe that the number of green patents has increased and that this increase is consistent when we vary the definition of what it means to be a "green" patent. In Figure 3b, we present a transformation of the data in the previous figure; rather than counts of patents and green patents, we present the fraction of patents filed by that startup cohort (by founding year) that are green patents. We observe an increase in the fraction over our sample, beginning at 15% in 1980 and increasing to 25% by the end of our sample.

We observe very comparable trends in the trademark filings as we did in the patent filings. In Figure 4a, we present the total number of trademarks (on the left vertical axis) and green trademarks (on the right vertical axis) by startup founding cohort (on the horizontal axis). We find a similar increase in the filing of trademarks and greenmarks. In Figure 4b, we present the fraction of trademarks filing by each startup cohort that are green trademarks. We observe an increase from 2% at the beginning of the sample to 5% towards the end of the sample.

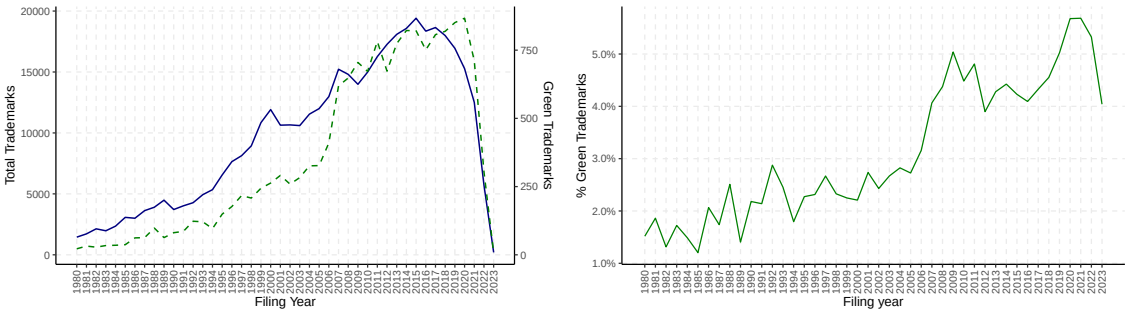


(a) Patent filings by year

(b) Proportion of green patent filings by year

Figure 3: Patent Filing Activity by Startups

Notes: The horizontal axis denotes filing year. Panel (a) presents counts of patents filed by startups in our sample. The blue line represents the annual number of patent filings and is associated with the scale on the left vertical axis. The green lines represent the annual number of green patent filings and is associated with the scale on the right vertical axis. Panel (b) presents the annual proportion of green patents filed by startups. Green patents follow the definitions outlined in Section 2.3.



(a) Trademark filings by year

(b) Proportion of green trademark filings by year

Figure 4: Trademark Filing Activity by Startups

Notes: The horizontal axis denotes filing year. In Panel (a) the blue line represents the annual number of trademark filings in that year by startups in our sample and is associated with the scale on the left vertical axis. The green line represents the total number of green trademark filings and is associated with the scale on the right vertical axis. In Panel (b) the green line represents the annual proportion of green trademark filings in that year for startups in our sample . Green trademarks ("greenmarks") follow the definition outlined in Section 2.3.

4 Green Startups and Their Green Innovations

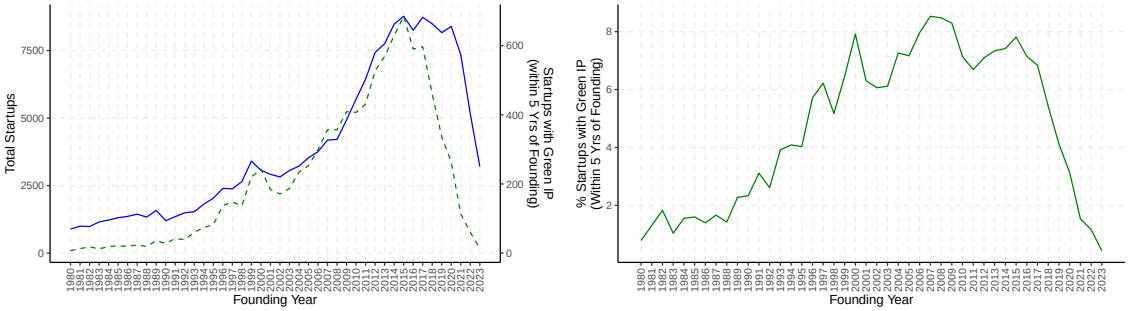
In this section, we attempt to classify some startups as green startups and compare their green innovative activities to non-green startups'. An immediate concern is how to classify startups as green versus non-green. We adopt two definitions. For the first definition "green-innovation startups", we classify startups based on startups' records in green innovations and products. We classify startups who engage in green innovation early in their lives as green startups. This innovation can be either patent or trademark filings and must occur within the first five years of the startup's life.

For the second definition of "self-labeled green startups", we utilize business descriptions of startups and quantify the green keyword ratio of the business description. Startups with high ratios of green keywords are classified as green and others are non-green. The business description is written by the startup, hence why we refer to this as self-labeled green startups.

4.1 Green-innovation Startups

Green-innovation startups are defined as those startups which file a green patent or trademark within the first five years after founding. This classification of startups is based on observable innovation outcomes.

Figure 5a presents the time series of green-innovation (right axis) and total startup formations (left axis). The plot reveals that the founding of green-innovation startups steadily increased in parallel with overall startup founding from near-zero levels in the 1980s. By the mid-1990s, the number of green-innovation startups accelerates, likely due to the increased global awareness of the Kyoto Protocol (1997) that raised the profile of sustainable technologies and green patenting (Miyamoto and Takeuchi 2019). Figure 5b presents the time series of the proportion of startups founded each year that are green-innovation startups. The percentage of green-innovation startups increases from around 2% in the 1990s to 8% in the early 2000s and has remained stable. Note that there are two layers of data truncation present in Figure 5. First, the lag in data collection by Pitchbook is visible starting in 2021 as shown by the solid blue line representing total startups in Figure 5a. Second, the delay in public disclosure of patent filings begins in 2019 in Figure 5a by the dashed green line. This is likely due to the potentially lengthy patent prosecution (and trademarking) time. Therefore, we are under-counting green-innovation startups in the latest period end of our sample.



(a) Annual Startup Formation

(b) % of Startups with Green Patents or Trademarks within 5 Years of Founding

Figure 5: Green-innovation Startup Formation Activity

Notes: Panel (a) presents the count of startups founded each year. The blue line corresponds to all startups and is associated with the vertical axis on the left. The green dashed line corresponds to startups with any green innovations within 5 years of founding and is associated with the vertical axis on the right. Panel (b) presents the proportion of startups with any green innovations within 5 years of founding founded each year.

4.1.1 Spatial distribution of green-innovation startups

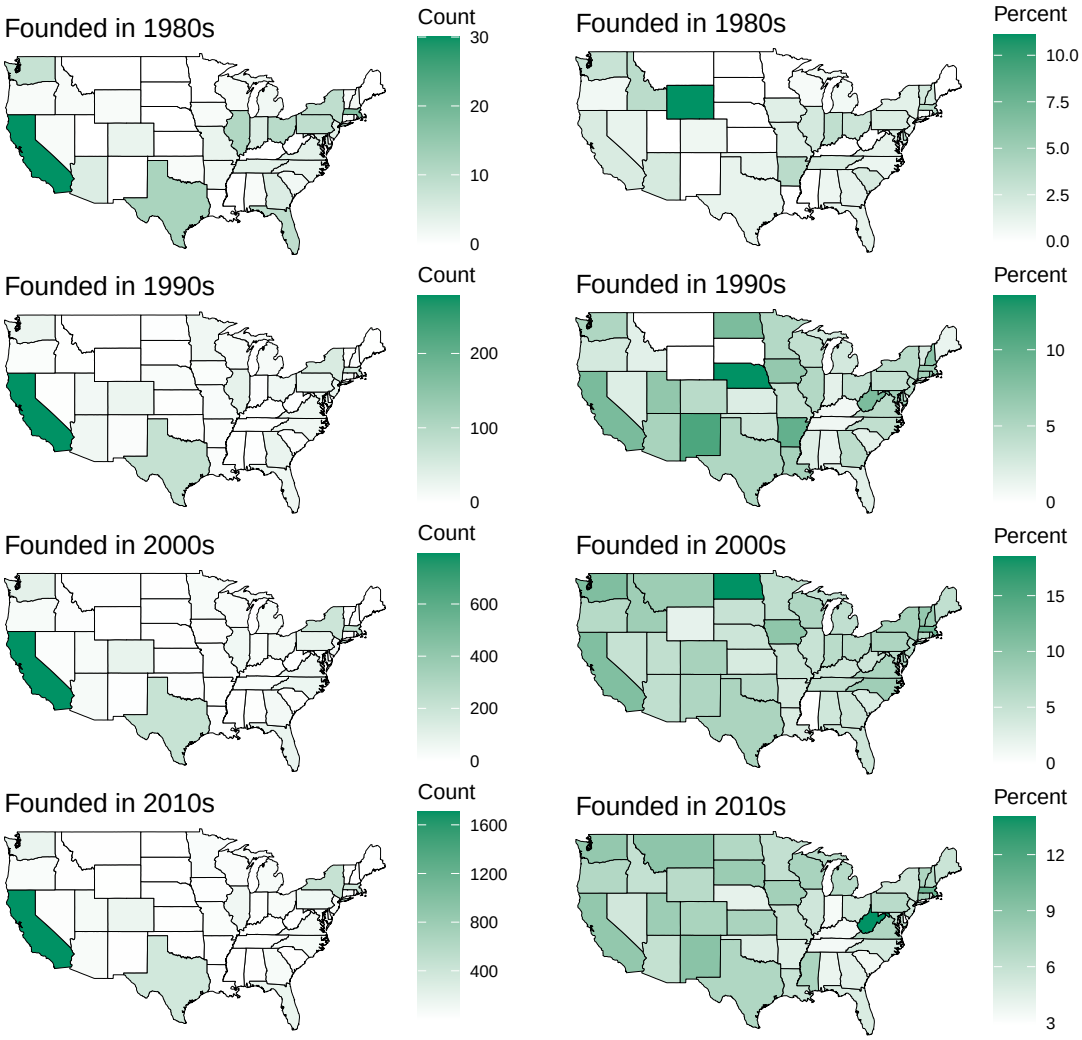
We next turn to the spatial distribution of green-innovation startups. In Figure 6a we present *counts* of green startups by state and decade of founding. California, Massachusetts, Texas, and New York have among the highest concentrations of green-innovation startups across all decades. Given the agglomeration effects of technology and venture-capital investment hubs, these states persistently attract entrepreneurs and are therefore the primary spawning points for startups, whether green or non-green. In the 1980s, California, Massachusetts, Texas, and New York in aggregate are where 39% of all green-innovation startups are spawned. In 1990s, over 50% of green-innovation startups were founded in these states. In the 2000s and 2010s, this number has increased to 55% and 60%, respectively.

However, the relative concentration of green-innovation startups compared to all startups founded by state paints a different picture. In Figure 6b we present green startup *proportions* by state and decade of founding. In the 1980s, Wyoming had the highest percentage of green-innovation startups founded at 11.1%. For the rest of the states, this number ranges from 0.5 to 3.8%. In the 1990s, Nebraska and New Mexico had green-innovation startup founding percentages exceeding 10%. 12 out of the 50 states plus Washington D.C. have green-innovation startup founding percentages between 5 and 10%. These 12 states do not include California, Texas, New York, or Massachusetts. In the 2000s, the state-by-state percentage of green-innovation startups spawned in North Dakota, Massachusetts, Washington, and California exceed 10%. 33 out of 50 states plus D.C.

have percentages of green-innovation startups exceeding 5%. In the 2010s, that number increases to 36. Overall, we see an overall increase in state-by-state green-innovation startup founding rates with no clear leader across all decades.

Restricting attention to the larger states, the dominant pattern in this figure is one of substantial growth in both the absolute number and the relative share of green-innovation startup foundings over time, particularly from the 1980s through the 2000s. California illustrates this pattern most clearly: the number of green-innovation startups increases from 30 in the 1980s to 278 in the 1990s, 792 in the 2000s, and 1,703 in the 2010s, while the corresponding green share rises from 2.0% to 8.4%, peaks at 10.9% in the 2000s, and then moderates to 8.6% in the 2010s. Massachusetts exhibits a similar trajectory, with green-innovation startup counts increasing from 17 to 58, 221, and 390, alongside an increase in the green share from 3.5% in the 1980s to 12.3% in the 2000s, before easing to 10.6% in the 2010s. Washington follows the same broad pattern, rising from 8 green-innovation startups in the 1980s to 99 in the 2000s and 172 in the 2010s, while the green share increases from 2.9% to 11.0% and then settles at 8.5%. More generally, the evidence suggests that, among major states, this category of green entrepreneurship moved from being a highly marginal component of startup activity in the 1980s to a much more economically meaningful presence by the 2000s and 2010s.

At the same time, the figure indicates that for many large states the relative importance of green-innovation startup formation peaks in the 2000s rather than continuing to rise monotonically into the 2010s. New York, for example, increases from 9 green-innovation startups in the 1980s to 463 in the 2010s, while its green share rises from 1.5% to 7.8% in the 2000s and then declines to 5.6% in the 2010s. Texas follows a similar pattern, with counts increasing from 12 to 372, but the share rising from 1.2% in the 1980s to 7.3% in the 2000s before slipping slightly to 7.0% in the 2010s. Pennsylvania rises from 1.9% in the 1980s to 7.1% in the 2000s and then declines modestly to 6.6%, while North Carolina increases from 1.0% to 8.3% and then falls back to 5.1%. Florida also remains below the leading states, moving from 1.2% in the 1980s to 4.3% in the 2000s and 4.9% in the 2010s. Overall, once smaller states are set aside, the evidence points to a common pattern in which larger states experience a marked increase in the level of green startup activity over time, together with a substantial rise in green startup intensity through the 1990s and 2000s, followed in many cases by a partial retrenchment in the 2010s.



(a) Count (b) Proportion

Figure 6: Number and Proportion of Green-Innovation Startups by State

Notes: Panel (a) presents the number of startups founded with green innovations within 5 years by state for each decade. Panel (b) presents the number of startups founded with green innovations within 5 years divided by total number of startups founded by state for each decade. Green innovation is defined as having filed either a a green trademark or patent within the first 5 years post-founding.

4.1.2 Sectoral Distribution of Green-Innovation Startups

In addition to full-sample startup counts, Table 3 also presents counts and proportions of green innovation startups founded in each decade of our sample. For many industries, the 2000s featured the highest proportion of green innovation startups. This applies to the Chemicals and Gases industry (26.6% in the 2000s versus a full sample proportion of 14.9%), Commerial Products (11.8% versus a full sample proportion of 6.5%), and Energy Equipment (42.9% versus a full sample proportion of 30.8%). The 2000s appear to be a decade when green innovation was central within many industries. The 2010s feature increased *counts* of green-innovation startups, but with declining within-industry *proportions*. The Energy Equipment industry is both a large industry in terms of green-innovation startup counts as well as in terms of green-innovation startup proportion. Unlike the Energy Equipment industry, the Software industry is large in terms of counts of green-innovation startups but with a low proportion (7.6% in the 2010s). The Commercial Banks and Insurance industries feature single digit numbers of the proportion of green-innovation startups. A similar pattern is found in the Restaurants, Hotels and Leisure industry (1.2%), Retail (2.5%) and Apparel and Accessories (2.3%) industries.

4.2 Self-labeled Green Startups

We next consider self-labeled green startups, those whose business descriptions have a high proportion of green keyword intensity. We consider startups whose green keyword intensity lies in the upper quartile of the sample (above the 75^{th} percentile cutoff) to be green startups. In Figure 7a we present the time series of the number of startup formation along with the annual proportion of green startups. During the 1980s, an average of 277 green startups were founded each year. The annual number increased throughout the sample and exceeded 2000 per year in the late 2010s.[21] In Figure 7b we present the annual proportion of green startups among all startups founded. During the 1980s, 25% of startups were green. This proportion declined during the 2000s to less than 20% and increased again in the 2010s. Towards the end of our sample, in the late 2010s, again roughly 25% of startups were green.

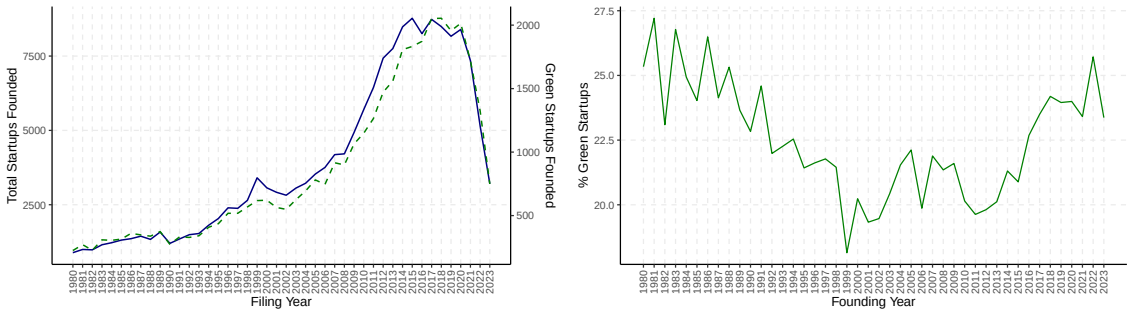
The classification of startups into self-labeled green versus non-green is based on self-reported business descriptions. To validate the measure we first focus on the differences in green innovations

21. We attribute the decline in the post-2020 sample to truncation biases in the data gathering process of Pitchbook and Revelio. As seen in Figure 7b, this truncation bias does not appear to systematically select for green versus non-green firms.

Table 3: Number and Proportion of Green-innovation Startups by Founding Decade and Industry Group

Industry Group	All Year		1980s		1990s		2000s		2010s	
	Count	(%)	Count	(%)	Count	(%)	Count	(%)	Count	(%)
Agriculture	69	(8.5)	0	(0.0)	1	(1.6)	8	(7.7)	14	(3.8)
Apparel and Accessories	56	(2.3)	0	(0.0)	2	(1.1)	7	(1.9)	12	(1.2)
Capital Markets/Institutions	32	(1.9)	4	(2.3)	2	(0.7)	13	(3.2)	20	(1.9)
Chemicals and Gases	168	(14.9)	6	(4.7)	21	(15.1)	55	(26.6)	91	(11.4)
Commercial Banks	5	(0.7)	1	(1.6)	0	(0.0)	2	(0.9)	3	(0.5)
Commercial Products	763	(6.5)	43	(2.6)	96	(6.4)	213	(11.8)	401	(4.3)
Commercial Services	1,097	(3.9)	32	(1.2)	149	(3.5)	325	(5.3)	569	(3.2)
Commercial Transportation	90	(7.1)	2	(1.2)	8	(4.5)	18	(8.2)	33	(3.4)
Communications and Networking	192	(8.0)	3	(1.6)	60	(9.0)	67	(9.1)	136	(7.4)
Computer Hardware	455	(13.9)	8	(3.8)	35	(11.8)	107	(16.9)	166	(11.4)
Construction (Non-Wood)	14	(7.7)	0	(0.0)	2	(7.7)	3	(10.3)	5	(3.6)
Consumer Durables	321	(7.4)	5	(1.9)	16	(4.3)	61	(9.6)	99	(4.9)
Consumer Non-Durables	159	(2.0)	3	(0.8)	7	(1.3)	30	(2.9)	49	(1.6)
Containers and Packaging	48	(6.1)	1	(1.1)	6	(5.3)	15	(13.9)	24	(3.7)
Energy Equipment	341	(30.8)	4	(7.5)	12	(16.9)	156	(42.9)	186	(30.6)
Energy Services	218	(15.8)	0	(0.0)	12	(9.7)	77	(21.2)	97	(13.4)
Exploration, Production and Refining	105	(11.2)	0	(0.0)	4	(5.6)	51	(16.2)	59	(10.6)
Forestry	9	(5.2)	0	(0.0)	1	(5.0)	5	(22.7)	7	(5.0)
Healthcare Devices and Supplies	308	(5.3)	6	(2.4)	46	(8.1)	95	(6.9)	156	(5.9)
Healthcare Services	125	(1.3)	3	(0.3)	16	(1.0)	34	(1.5)	59	(1.0)
Healthcare Technology Systems	222	(4.9)	0	(0.0)	14	(4.9)	63	(8.7)	85	(6.5)
IT Services	127	(3.1)	4	(1.3)	32	(3.6)	47	(3.6)	88	(3.2)
Insurance	20	(0.7)	0	(0.0)	2	(0.5)	3	(0.6)	5	(0.2)
Media	136	(2.4)	0	(0.0)	13	(3.0)	42	(3.8)	63	(2.6)
Metals, Minerals and Mining	23	(6.6)	1	(3.4)	1	(2.4)	6	(9.7)	9	(3.6)
Other Business Products and Services	28	(3.1)	0	(0.0)	1	(0.9)	8	(4.1)	13	(2.0)
Other Consumer Products and Services	0	(0.0)	0	(0.0)	0	(0.0)	0	(-)	0	(0.0)
Other Energy	13	(10.2)	0	(0.0)	1	(11.1)	7	(21.9)	8	(11.3)
Other Financial Services	105	(2.8)	1	(0.5)	17	(4.0)	22	(3.3)	46	(2.8)
Other Healthcare	1	(5.0)	0	(0.0)	0	(0.0)	0	(0.0)	0	(0.0)
Other Information Technology	1	(6.7)	0	(-)	0	(-)	0	(0.0)	1	(20.0)
Other Materials	55	(21.1)	0	(0.0)	7	(28.0)	11	(23.4)	20	(17.2)
Pharmaceuticals and Biotechnology	803	(10.8)	22	(9.1)	101	(15.7)	229	(13.5)	385	(13.1)
Restaurants, Hotels and Leisure	38	(1.2)	0	(0.0)	7	(2.0)	9	(1.6)	17	(1.0)
Retail	54	(2.5)	1	(0.6)	7	(2.8)	17	(4.6)	28	(2.2)
Semiconductors	171	(21.7)	11	(18.0)	25	(19.5)	63	(24.5)	105	(20.6)
Services (Non-Financial)	128	(1.9)	1	(0.2)	8	(1.3)	24	(2.2)	39	(1.2)
Software	2,365	(5.5)	11	(1.3)	181	(6.8)	615	(9.2)	922	(7.6)
Textiles	8	(5.0)	0	(0.0)	2	(9.1)	1	(5.0)	3	(2.6)
Transportation	156	(7.6)	2	(1.0)	7	(3.2)	35	(11.1)	46	(3.7)
Utilities	41	(12.9)	1	(9.1)	4	(12.1)	14	(24.1)	21	(9.7)
Total	9,070	(5.2)	176	(1.6)	926	(5.0)	2,558	(7.7)	4,090	(4.7)

Notes: Green innovation startups are defined as those startups which filed any green innovations (an eventually granted green patent or eventually registered green trademark) within 5 Years of founding. Percentage shares are in parentheses.



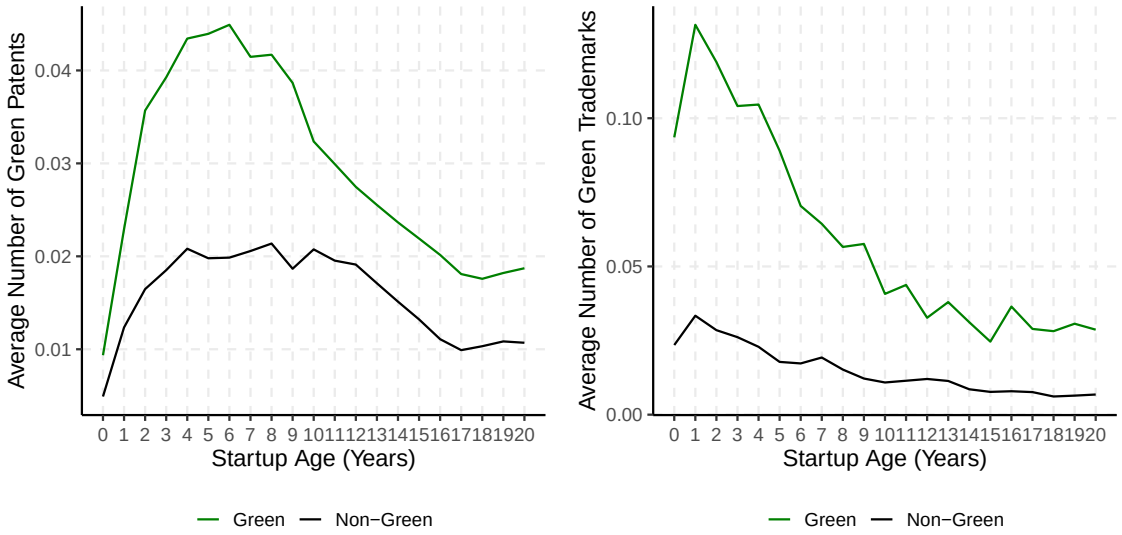
(a) Annual Startup Formation (b) % of Self-labeled Green Startups

Figure 7: Self-labeled Green Startup Formation Activity

Notes: Panel (a) presents the count of startups founded each year. The blue line corresponds to all startups and is associated with the vertical axis on the left. The green dashed line corresponds to green startups and is associated with the vertical axis on the right. Panel (b) presents the proportion of green startups founded each year. Green startups are based on the "self-labeled" classification and are those whose green keyword ratio is in the top quartile of the sample.

between green and non-green startups to establish that this classification is meaningful. In Figure 8, we plot the average number of green innovations filed over the life cycle of a startup. Here, the horizontal axis denotes startup age, rather than cohort or filing year. The left panel is patent filing activity and the right panel is trademark filing activity. We find that, green startups file more green patents and trademarks, on average, than non-green startups.

To understand how startups' life cycle is related to their green innovations, Figure 9 presents the arrival time of the startup's first green patent. The left panel is the cumulative distribution of that arrival time, meaning the probability that a startup files a green patent within T years of founding. The right panel plots hazard rates, the probability of filing the first green patent in year T, conditional on having never filed a green patent prior to year T. We observe that, for non-green startups, the overall probability of ever filing a green patent is lower, and these filings occur very early in their life cycle. For green startups, we observe something different. The overall probability of filing a green patent is higher. Moreover, although the likelihood of filing a green patent is high for young green startups, this probability remains high throughout the entire life. As one point of reference, a green startup's likelihood of filing their first green patent in year 2 since inception is as high as the likelihood for a non-green startup in year 6. Furthermore, green startups are twice as likely to have filed a green patent within the first three years as a non-green startup. Taken together, we infer that green startups file green patents earlier than non-green startups and are also more likely to continue pursuing green technologies (proxied by patent filing activity) than a non-green startup.



(a) Green Patent Filing Activity

(b) Green Trademark Filing Activity

Figure 8: Green Innovation Filing Activity by Green and Non-Green Startups

Notes: The horizontal axis corresponds to startup age. The vertical axis denotes the number of filings that startup files each year. Panel (a) presents annual filings of Green Patents. Panel (b) presents annual filings of Green Trademarks. The green line corresponds to green startups and the black line corresponds to non-green startups. Green patents and green trademarks follow the definitions outlined in Section 3. Green startups are based on the "self-labeled" classification and are those whose green keyword ratio is in the top quartile of the sample.

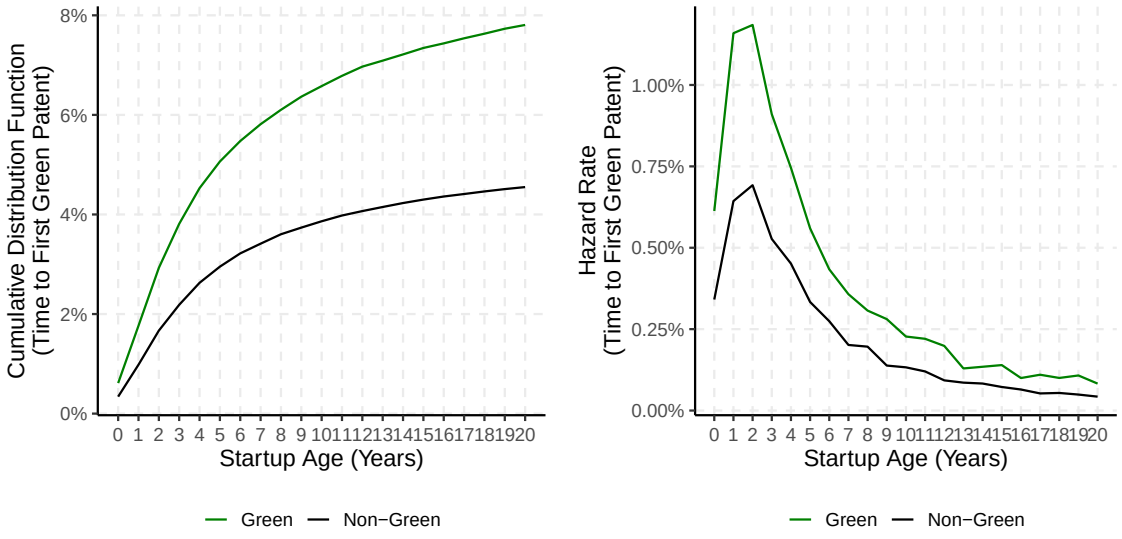


Figure 9: Startup Age in Filing of the First Green Patent

Notes: The horizontal axis corresponds to startup age. The vertical axis denotes the probability that a startup has filed a green patent. The left panel presents the cumulative probability that a start has filed a green patent within T years of founding, while the right panel presents the hazard rate that, conditional on not having filed a green patent by age $T-1$, the startup files a green patent in the next year.

In Figure 10, we find a similar story holds for green trademarking activity. Indeed, the relative magnitudes between green and non-green startups are more striking. The chance for a green startup to file its first green trademark in year 3 since inception is quadruple of that for a non-green startup. Consistent with this, in the right panel of the figure we observe that the hazard rate for first green trademark is much higher for green startups than non-green startups. A green startup is over three times as likely to have filed a green trademark in the first year as a non-green startup.

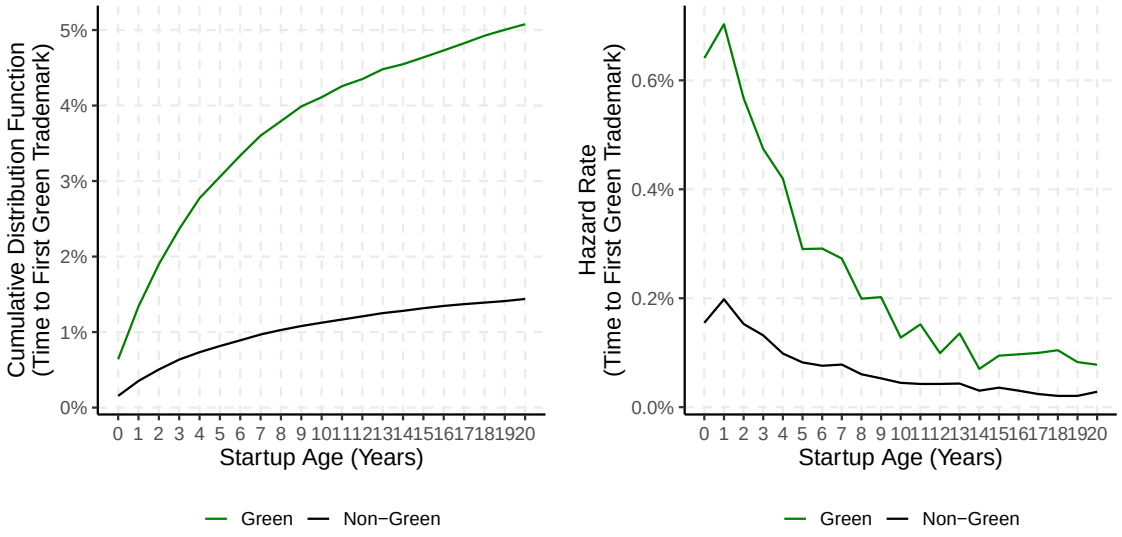


Figure 10: Startup Age in Filing of the First Green Trademark

Notes: The horizontal axis corresponds to startup age. The vertical axis denotes the probability that a startup has filed a greenmark. The left panel presents the cumulative probability that a start has filed a greenmark within T years of founding, while the right panel presents the hazard rate that, conditional on not having filed a greenmark by age $T-1$, the startup files a greenmark in the next year.

Overall, our analyses of the incidence and frequency of green innovations suggest that self-labeled green startups indeed create more green patents and trademarks than other non-green startups. This confirm that those entrepreneurs' self-disclosure is reasonable credible. In other words, they indeed walk their talk in terms of environment-oriented business activity.

4.2.1 Spatial distribution of self-labeled green startups

We next turn to the spatial distribution of green startup activity. In Figure 11a, we present counts of self-labeled green startups by state and decade. We observe that green startups are persistently formed in the same subset of states, which can be attributed to the stable relative economy size across states. California, Massachusetts, Texas, and New York rank among the highest concentrations of self-labeled green startups across all decades. In the 1980s, California,

Massachusetts, Texas, and New York in aggregate are where 30% of all self-labeled green startups are spawned. This number is very close to the number computed for the fraction of green-innovation startups for the same states. In 1990s, over 34% of self-labeled green startups are founded in these states. In the 2000s and 2010s, this number has increased to 40.3% and 49.8%, respectively. This trend parallels what we observed under the green-innovation criteria discussed above.

In Figure 11b, we present the percentages of self-labeled green startups by state. Similar to what we find under the green-innovation classification of startups, the relative intensity self-labeled green startups founding is quite dispersed across states and across time. In the 1980s, West Virginia has the highest percentage of self-labeled green startups founded at 53.3%. Oklahoma and Maine rank the second and third with around 40%. The other states' percentages are closer to 25%, which is the threshold we used to define a self-labeled green startup. In the 1990s, Oklahoma and West Virginia are the top two states with the highest percentage of self-labeled green startups at over 37%. This decrease indicates that the founding of self-labeled green startups perhaps become more dispersed across different states. In the 2000s, West Virginia and Maine lead the founding intensity of self-labeled green startups with percentages above 40%. In the 2010s, Maine become the leader at 39%. Some states increase rapidly in the percentage of self-labeled green startups across time: for instance, Wyoming surges from 11.1% in the 1980s to 29.6% in the 2010s. New Mexico is also a notable state that experienced an increase from 17.1% in the 1980s to 30.3% in the 2010s.

Restricting attention to the larger states, where the reported percentages are less likely to be driven by small-sample variation, the dominant pattern is that the absolute number of green startup foundings increases sharply over time, whereas the share of startup activity accounted for by green ventures often remains flat or declines modestly. California provides the clearest illustration: the number of green startups rises from 329 in the 1980s to 3,894 in the 2010s, while the corresponding green share declines from 22.0% to 19.6%, after reaching a trough of 17.6% in the 2000s. New York exhibits a similar pattern, with green startup counts increasing from 130 to 1,600, even as the green share moves slightly downward from 21.3% to 19.2%. Texas likewise expands from 256 to 1,245 green startups, while its green share declines from 25.7% to 23.5%. Comparable dynamics are evident in Florida, where green startup counts rise from 166 to 675 as the green share falls from 23.0% to 20.0%, and in Massachusetts, where counts increase from 118 to 844 while the corresponding share edges down from 24.5% to 23.0%. Collectively, these patterns indicate that the 2010s are principally characterized by an expansion in the scale of green entrepreneurial activity, rather than by an increase in its relative importance within overall startup formation.

A second notable feature of the data is that several large and economically important states exhibit a meaningful decline in green startup intensity over the sample period. Illinois declines from 23.9% in the 1980s to 20.2% in the 2010s; Pennsylvania from 27.6% to 23.7%; Washington from 24.7% to 21.4%; Michigan from 28.6% to 23.6%; and Maryland from 30.8% to 22.5%, despite substantial increases in the absolute number of green startups in each case. By contrast, a smaller group of medium-to-large states appears more stable over time: Arizona remains broadly flat, moving from 23.4% in the 1980s to 22.8% in the 2010s, while Georgia changes only modestly from 21.0% to 19.2%. North Carolina remains relatively elevated throughout the sample, though it too declines from 28.3% to 23.6%, and Colorado remains comparatively high, moving from 28.9% to 26.5%. Overall, once smaller states are excluded, the evidence points to a common pattern in which major states experience substantial growth in the level of green startup formation, but generally do not experience a commensurate increase in the proportion of startup activity that is green.

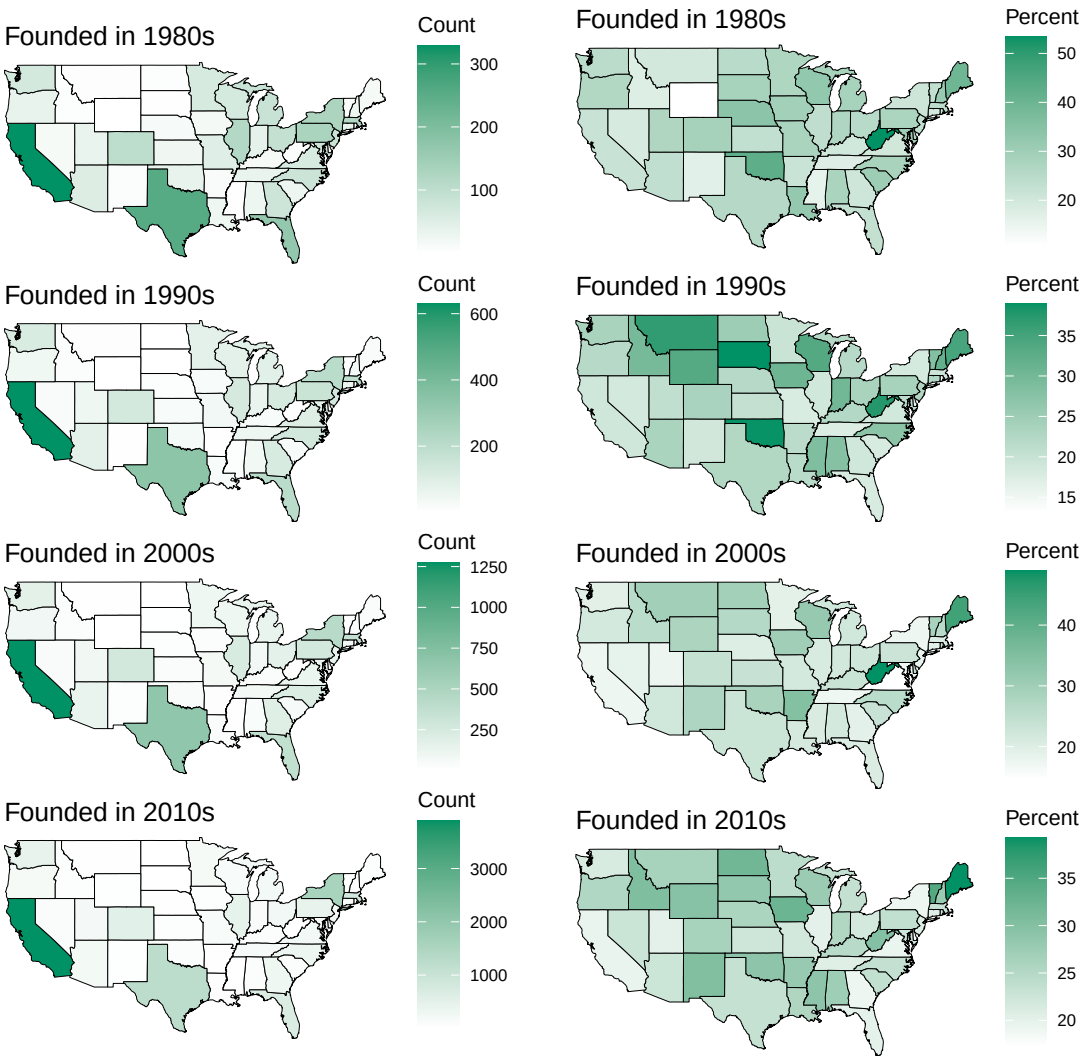
Taken together, we see that self-labeled green startups have a similar geographic distribution as overall trademarking and patenting activities, but that the tendency to cluster is lower than for startups as a whole. The relative proportion of startups which are self-labeled as green is higher in states with lower rates of startup formation.

4.2.2 Sectoral distribution of self-labeled green startups

We now examine sector-specific counts of green startup formation, and the calculated proportion of green startups in Table 4. We find that the sectoral distribution of green startups is distinct from that of general startups as shown in Table 2. Although Software is the most represented industry among all startups in our sample (43,383 startups reported in Table 2), only 7.85% of those startups are green. In addition, the Healthcare Services industry is the fourth largest in terms of overall startup counts, but only 2.94% of these startups are green. On the other hand, there are 788 semiconductor startups in our sample, and 35.91% of them are green. The industry of the second highest ratio of green startups is utilities: 29.15% of 319 utility startups in our sample are green. These differences likely reflect inherent differences between the sectoral composition of green startups and startups overall.

In addition to full-sample startup counts, Table 4 also presents counts and proportions of green startups founded each decade of our sample. The industries fall into two categories.[22] There are high green-proportion industries with proportions between 50 and 70 percent. Industries in this

22. Excluding industries with very few startups in our sample.



(a) Number (b) Proportion

Figure 11: Number and Proportion of Self-labeled Green Startups by State

Notes: Panel (a) presents the number of self-labeled green startups by state for each decade. Panel (b) presents the number of self-labeled green startups divided by total number of startups founded by state for each decade. Green startups are those in the top 25% Green keyword ratio based on the full sample of startups.

category include Agriculture, Exploration, Production, and Refining, Energy Services, Chemicals and Gases, and Containers and Packaging. The other category consists of low green-proportion industries, those with green proportions between 10 and 30 percent. This includes industries such as Software, IT Services, Communications and Networking, Media, Restaurants, Insurance, and Healthcare Services. Most financial service industries are contained in the low-proportion group. Furthermore, the largest industries are not the greenest industries. Commercial Products, Commercial Services, and Software are large industries by counts of green startups, but these industries are in the low proportion group. Finally, we observe that, though the overall green proportion is not meaningfully different between the 1980s and 2010s, there are several sectors which become more green-intensive over time. Energy Equipment increases from 54.7% in the 1980s to 64.5% in the 2010s, with an increase of 9.8 percentage points. Among the low-proportion industries, Consumer Non-Durables increases from 20.2% in the 1980s to 24.1% in the 2010s, with an increase of 3.9 percentage points. Over the sample, Energy Services declines in green-proportion from 62.3% to 56.4%, with a decline of 5.9 percentage points. Among the low-proportion industries, Software declines from 17.0% to 13.9%, with a decline of 3.1 percentage points. Other industries with declines of similar magnitude incldue Containers and Packaging (3.9%), Communications and Networking (3.0%), and Non-Financial Services (2.9%).

Table 4: Number and Proportion of Self-labeled Green Startups by Founding Decade and by Industry Group

Industry Group	All Years		1980s		1990s		2000s		2010s	
	Count	(%)	Count	(%)	Count	(%)	Count	(%)	Count	(%)
Agriculture	534	(66.0)	29	(63.0)	33	(53.2)	64	(61.5)	213	(58.2)
Apparel and Accessories	445	(18.2)	20	(17.1)	31	(17.7)	56	(14.8)	159	(15.4)
Capital Markets/Institutions	254	(15.4)	18	(10.4)	36	(13.0)	54	(13.1)	131	(12.6)
Chemicals and Gases	659	(58.3)	61	(47.7)	72	(51.8)	114	(55.1)	417	(52.3)
Commercial Banks	39	(5.4)	3	(4.8)	4	(5.0)	11	(5.1)	31	(4.8)
Commercial Products	4,484	(38.4)	591	(36.3)	512	(34.0)	708	(39.4)	3,416	(36.6)
Commercial Services	7,467	(26.6)	807	(29.7)	1,123	(26.6)	1,434	(23.2)	4,805	(27.1)
Commercial Transportation	498	(39.1)	61	(35.7)	63	(35.6)	89	(40.6)	357	(36.3)
Communications and Networking	304	(12.6)	27	(14.5)	50	(7.5)	89	(12.1)	213	(11.5)
Computer Hardware	1,013	(31.0)	61	(28.6)	73	(24.6)	136	(21.5)	378	(25.9)
Construction (Non-Wood)	82	(44.8)	2	(25.0)	6	(23.1)	11	(37.9)	50	(36.5)
Consumer Durables	1,135	(26.2)	65	(24.2)	77	(20.6)	159	(25.0)	485	(24.1)
Consumer Non-Durables	2,222	(27.5)	76	(20.2)	128	(24.0)	276	(27.1)	727	(24.1)
Containers and Packaging	426	(54.3)	52	(56.5)	57	(50.4)	65	(60.2)	337	(52.6)
Energy Equipment	758	(68.4)	29	(54.7)	44	(62.0)	236	(64.8)	392	(64.5)
Energy Services	860	(62.4)	43	(62.3)	66	(53.2)	206	(56.7)	408	(56.4)
Exploration, Production and Refining	596	(63.3)	40	(60.6)	37	(52.1)	189	(60.2)	333	(59.6)
Forestry	88	(51.2)	13	(52.0)	13	(65.0)	11	(50.0)	70	(49.6)
Healthcare Devices and Supplies	1,069	(18.5)	50	(20.1)	98	(17.3)	214	(15.5)	469	(17.7)
Healthcare Services	1,287	(13.0)	111	(12.5)	186	(11.8)	277	(12.0)	744	(12.2)
Healthcare Technology Systems	779	(17.3)	12	(12.5)	46	(16.1)	133	(18.3)	226	(17.2)
IT Services	493	(12.2)	27	(9.0)	110	(12.5)	134	(10.3)	295	(10.7)
Insurance	308	(10.7)	31	(7.4)	32	(8.1)	66	(13.9)	209	(9.6)
Media	685	(12.1)	26	(14.2)	61	(14.0)	109	(9.8)	278	(11.5)
Metals, Minerals and Mining	137	(39.6)	16	(55.2)	13	(31.0)	26	(41.9)	94	(37.8)
Other Business Products and Services	170	(18.8)	17	(16.3)	23	(20.4)	29	(15.0)	108	(16.4)
Other Consumer Products and Services	5	(33.3)	0	(0.0)	1	(100.0)	0	(-)	1	(33.3)
Other Energy	71	(55.9)	2	(40.0)	7	(77.8)	17	(53.1)	39	(54.9)
Other Financial Services	602	(15.8)	29	(13.6)	58	(13.7)	85	(12.7)	229	(13.8)
Other Healthcare	4	(20.0)	0	(0.0)	1	(50.0)	0	(0.0)	1	(25.0)
Other Information Technology	2	(13.3)	0	(-)	0	(-)	2	(66.7)	2	(40.0)
Other Materials	149	(57.1)	2	(18.2)	11	(44.0)	24	(51.1)	52	(44.8)
Pharmaceuticals and Biotechnology	2,083	(27.9)	64	(26.6)	188	(29.3)	434	(25.5)	809	(27.5)
Restaurants, Hotels and Leisure	371	(11.5)	28	(12.1)	43	(12.0)	50	(8.7)	187	(10.8)
Retail	324	(15.0)	18	(11.5)	36	(14.6)	51	(13.7)	177	(13.7)
Semiconductors	196	(24.9)	13	(21.3)	26	(20.3)	61	(23.7)	120	(23.6)
Services (Non-Financial)	1,291	(19.3)	100	(24.0)	122	(19.6)	222	(20.0)	665	(21.1)
Software	6,681	(15.4)	147	(17.0)	391	(14.7)	889	(13.3)	1,681	(13.9)
Textiles	76	(47.2)	5	(38.5)	8	(36.4)	7	(35.0)	40	(34.8)
Transportation	580	(28.3)	59	(28.1)	77	(35.2)	69	(21.8)	329	(26.4)
Utilities	197	(61.8)	5	(45.5)	16	(48.5)	34	(58.6)	122	(56.2)
Total	39,424	(22.5)	2,760	(24.8)	3,979	(21.4)	6,841	(20.6)	19,799	(22.9)

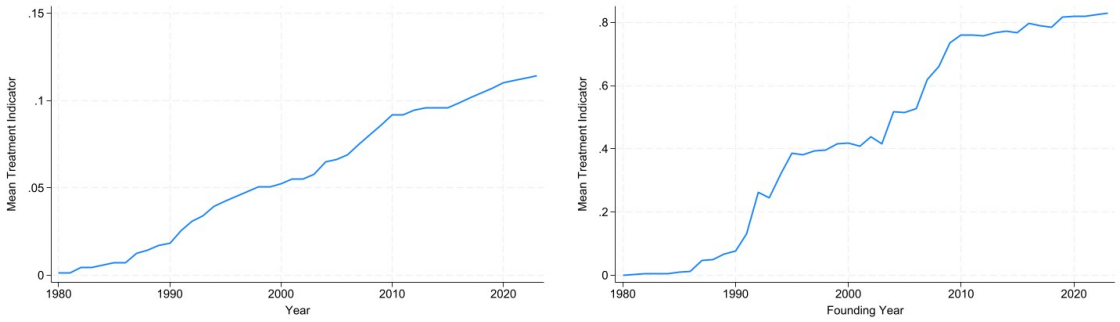
Notes: Green startups are those whose business description % Green Keywords Ratio is above the 75^{th} percentile of the full sample. All Years exceeds the sum of columns 1980s, 1990s, 2000s and 2010s due to the presence of startups startups founded prior to 1980 and after 2019 in our sample. Percentage shares are in parentheses.

5 Factors for Green Entrepreneurship and Innovations

5.1 Evidence from green research institutes

Universities play a vital role in innovation, and research institutes serve to agglomerate domain expertise and concentrate research infrastructure. They serve as a nexus for repeated interactions between scientists engaging in basic and applied research, students, and industry partners. For all these reasons, we expect that green-oriented research institutes serve to stimulate local green startup formation and innovation.

Figure 12a presents a time series of the proportion of commuting zones which have a green research institute. We consider a commuting zones as "treated" once the first local green research institute is founded. In total, the sample includes the opening of 272 green research institutes in 82 unique commuting zones. Figure 12b presents the same data, this time at the startup-level. The proportion of firms founded in green research institute commuting zones has trended upwards over time. By the late 2010s, over 80% of startups were founded in commuting zones with a local green research institute.



(a) Proportion of treated commuting zones. (b) Proportion of treated startups.

Figure 12: The Proportion of Treated Sample by Green Research Institute Formation

Notes: Panel (a) plots the time series of the proportion of commuting zones with a local green research institute. Panel (b) presents the time series of the proportion of startups founded each year in commuting zones with a local green research institute. Local green research institutes are those affiliated with a university in the same commuting zones.

We utilize a staggered difference-in-differences framework to study the response of innovation-related outcome variables to the founding of local green research institutes. We compare the increase in green innovations created by all startups in treated commuting zones to that in untreated commuting zones to estimate the role of local green research institutes. Using ordinary least squares

regression, we estimate the following model:

$$Y_{ict} = \alpha_G \text{Green Research Institute}_{c,t-1} + \text{Commuting Zone FE} + \text{Year FE} + \text{City FE} + \varepsilon_{ict}, \quad (1)$$

where Y_{ict} is the proportion of green patents or green trademarks filed by firm i which was founded in year t in commuting zones c within the first N years of firm life (thus, each observation corresponds to a unique startup); Green Research Institute$_{ct}$ is the treatment indicator which takes a value of 1 if there is a green research institute in commuting zones c founded in year t or earlier, meaning the firm was created contemporaneous to or after a green research institute was established.[23] To avoid issues related to truncation bias, we exclude startups founded after 2019, as that is the population for which intellectual property grants may have already been filed but yet to be rejected or assigned by the end of our sample.

In Table 5 we report results on startups' green innovations following the opening of a local green research institute. We find that early-stage green trademark filing activity by startups increases after a green research institute is opened in the same commuting zone. The increase of 0.7–1 percentage points is statistically significant over the first five years of startup life. This is consistent with our hypothesis that green research institutes serve to stimulate green commercialization. Notably, we do not find positive or significant results for green patents. In terms of economic magnitude, in Table 5, Column (4), the coefficient of Green Research Institute$_{ct}$ implies that when Green Research Institute$_{ct}$ equals to one, the green trademark ratio increases by 0.008, which corresponds to 0.008/0.039=20.5% relative to the unconditional mean green trademark ratio of 0.039.[24] Nevertheless, the creation of such institutes does not lead to additional green patents being filed by local startups.

Next we examine the aggregate effects in the treated commuting zones. Using Poisson regression, we estimate the following model:

$$Y_{ct} = \exp\left(\alpha_G \text{Green Research Institute}_{c,t-1} + \text{Commuting Zone FE} + \text{Year FE} + \varepsilon_{ct}\right), \quad (2)$$

where Y_{ct} is the count of green startup formation in commuting zone c in year t. Our primary object of interest is α_G, the coefficient on Green Research Institute.

23. Note that these year fixed effects absorb the usual "Post" coefficient.

24. Appendix Table B.3 summarizes the regression variables. In the OLS specification $Y = \alpha + \beta X + \varepsilon$, a one-unit increase in X changes the conditional expectation of Y by β units. When X is an indicator, moving from 0 to 1 implies $\Delta Y = \beta$.

Table 5: Startup-level response to the founding of a local green research institute.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Green Trademarks (ratio)						
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$	$t \to t+6$
Green Research Institute$_{c,t-1}$	0.004	0.004	0.007*	0.008**	0.009**	0.010***
	(0.006)	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)
Commuting Zone FE	Y	Y	Y	Y	Y	Y
Founding Year FE	Y	Y	Y	Y	Y	Y
Startup HQ FE	Y	Y	Y	Y	Y	Y
Observations	16,203	25,325	31,285	35,759	39,200	41,842
Adjusted R^2	0.0170	0.0220	0.0210	0.0170	0.0160	0.0190
Panel (b): Green Patents (ratio)						
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$	$t \to t+6$
Green Research Institute$_{c,t-1}$	0.019	0.000	0.013	0.009	0.014	0.017
	(0.022)	(0.019)	(0.015)	(0.016)	(0.018)	(0.018)
Commuting Zone Fixed Effects	Y	Y	Y	Y	Y	Y
Founding Year FE	Y	Y	Y	Y	Y	Y
Startup HQ FE	Y	Y	Y	Y	Y	Y
Observations	3,828	6,903	9,802	11,929	13,444	14,501
Adjusted R^2	−0.0100	0.0140	0.0190	0.0200	0.0190	0.0210

Notes: The table presents startup-level linear regression results for intellectual property-related outcomes following the founding of a local green research institute. The specification is given in Equation (1). Panel (a) presents results using the proportion of green trademarks as the outcome variable, while Panel (b) presents results using the proportion of green patents as the outcome variable. In Column (1) the outcome variable is the proportion measured over the first year of the firm (age 0). In Columns (2)–(6) the proportion is measured over the first 2, 3, … 5 years of the firm (ages 1 through 6). Green Research Institute$_{c,t-1}$ is the treatment indicator which takes a value of 1 if there is a green research institute in commuting zones c founded in year $t-1$ or earlier. Standard errors in parentheses are clustered at the commuting zone level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

In Table 6 we report results on green startup formation following the opening of a local green research institute. In our staggered difference-in-differences setup, we find that green startup formation in treated commuting zones increases by more than green startup formation in untreated commuting zones. Furthermore, the effect remains statistically significant across several methods of classifying green startups.

In Panel (a) we report results on annual frequency of green startup formation based on green innovation activities. In Column (1), we classify startups which file a green patent or trademark within the first year of founding to be green. In Columns (2)–(6) we use alternative classification rules and require that green startups file at least one green patent or trademark within the first n years of founding, with n ranging from 2 to 6 years. In the Poisson specification, the coefficient estimate on Green Research Institute$_{t-1}$ of 0.226 in Column (3) implies that commuting zones with green research institutes experience a $\exp(0.226) - 1 = 25.4\%$ increase in green-innovation startups over $(t, t+3)$. Using the unconditional mean of green-innovation startups (based on patent and trademark filings within the first three years post-founding) of 1.05,[25] this corresponds to an increase of approximately $1.05 \times (\exp(0.226) - 1) \approx 0.27$ additional startups per commuting-zone-year. These magnitudes are consistent with a local innovation ecosystem mechanism in which research institutes facilitate green commercialization through knowledge spillovers, talent concentration, and reduced coordination frictions.

In Panel (b) of Table 6, the classification of startups into green versus non-green is based on the green keyword ratio in the business description. In Column (1), we count the number of firms for which the intensity of green terms in the business description is above the full sample median. In Column (2), we use the 75$^{\text{th}}$ percentile. In Column (3), we use the 90$^{\text{th}}$ percentile. We see that the magnitude of the coefficient declines but the effect remains positive. In Columns (4)–(6), we use an industry-specific cutoff based on the startup's industry classification in Pitchbook. This is intended to capture differences across industries in the tendency to include climate-related terms in the business description. In Column (4) we see that the effect is statistically significant when we use the industry median intensity as the cutoff. In Column (5) we see the effect remains significant when we use the industry-specific 75$^{\text{th}}$ percentile. In Column (6) we use the industry-specific 90$^{\text{th}}$ percentile intensity as the cutoff and we see that the effect is positive but no longer statistically significant. The economic interpretation of the Poisson regression coefficient of 0.202 in Column (1) is that in a treated commuting zone (for which Green Research Institute$_{t-1}$ equal to one) self-

25. Appendix Table B.7 reports summary statistics for the regression variables.

labeled green startup formation increases by $\exp(0.202) - 1 = 22.4\%$. Using the unconditional mean of N^{50} (7),[26] this translates into an increase from 7 to 8.57 green startups; an additional 1.57 startups per commuting-zone-year. As a comparison, this increase is equal to 22.4% of the baseline formation rate among untreated counties and is economically meaningful. Furthermore, if persistent, the estimate implies roughly 8 additional green startups over a five-year horizon for a treated commuting zone.[27]

5.2 Evidence from green ARRA funding

The importance of government action to confront the challenges of climate change has been well-discussed in the literature. In this section, we study the impact of a big push for the green transition: the American Recovery and Reinvestment Act (ARRA) that allocated over $350 billion of direct government spending and an additional $260 billion in tax reductions (Aldy 2013). Furthermore, as emphasized by Popp et al. (2021), roughly 17% of direct spending was targeted at green investments. In Table 7 we report annual disbursements of green ARRA funds between 2009 and 2012. Over 60% of the green ARRA funds were spent in the first year after passage and essentially all funds were spent within the first three years.

We use data on the spatial allocation of green ARRA funding to study the impact of public funding on green innovations. Specifically, we use the heterogeneous distribution of funds across commuting zones. In total, roughly $51 billion was allocated to the EPA and DOE to allocate to green investment, with the bulk spent in 2009, the first year following passage of the legislation. Motivated by these institutional details, we use the total amount of ARRA funding given to a commuting zone as our measure of funding. We then scale this by population to get a per-capita measure of ARRA funding intensity.[28] Using ordinary least squares regression, we estimate the

26. Appendix Table B.5 presents summary statistics for the regression variables

27. In a Poisson regression with a log link, $\ln \mathbb{E}[Y \mid X] = \alpha + \beta X$. When $X = 1$, the conditional mean is $\mathbb{E}[Y \mid X = 1] = \exp(\alpha + \beta)$; when $X = 0$, it is $\mathbb{E}[Y \mid X = 0] = \exp(\alpha)$. Therefore,

$$\frac{\mathbb{E}[Y \mid X = 1]}{\mathbb{E}[Y \mid X = 0]} = \exp(\beta),$$

so switching X from 0 to 1 implies a proportional change in Y of $\exp(\beta) - 1$.

28. To avoid concerns about the endogenous reallocation of labor in response to the availability of ARRA funds, we scale by population as of 2000.

Table 6: Commuting zone-level response to the founding of a local green research institute.

Panel (a): Green-innovation Startups (counts)						
	(1)	(2)	(3)	(4)	(5)	(6)
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$	$t \to t+6$
Green Research Institute$_{c,t-1}$	0.217*	0.174*	0.226**	0.171*	0.170*	0.160*
	(0.111)	(0.090)	(0.090)	(0.088)	(0.091)	(0.096)
Commuting Zone FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Observations	3,735	4,016	4,133	4,220	4,234	4,252
Pseudo-R^2	0.492	0.596	0.658	0.686	0.696	0.703
Panel (b): Self-labeled Green Startups (counts)						
	N^{50}	N^{75}	N^{90}	N^{50}_{ind}	N^{75}_{ind}	N^{90}_{ind}
Green Research Institute$_{c,t-1}$	0.202**	0.160**	0.096	0.203**	0.154*	0.103
	(0.088)	(0.074)	(0.081)	(0.093)	(0.082)	(0.085)
Commuting Zone FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Observations	9,258	9,199	9,050	9,238	9,149	8,925
Pseudo-R^2	0.837	0.775	0.711	0.848	0.794	0.731

Notes: The table presents county-level Poisson regression results for green startup formation rates following the founding of a local green research institute. The specification is given in Equation (2). In Panel (a) the classification of green startups is based on the startup's green patent and trademark filings. In Column (1), we consider a strict definition of green innovation startups which requires that the firm file a green patent or trademark within one year of founding. In Columns (2)–(6) we use a sequence of classification rules based on whether the startup filed a green patent or trademark within n years of foudning, with n ranging from 2 to 6. In Panel (b) the classification of green startups is done based on the self-labeling in the business description. In Column (1), the outcome variable is the count of new green startups founded each year, where green startups are those whose green keyword intensity is above the full-sample median. In Column (2), we use the 75^{th} percentile of green keyword intensity as the cutoff. In Column (3), we use the 90^{th} percentile of green keyword intensity as the cutoff. In Columns (4)–(6), we repeat the same sequence of cutoffs, but using industry-specific cutoffs rather than full sample cutoffs. Green Research Institute$_{c,t-1}$ is an indicator value equal to 1 when commuting zone c has a local green research institute by year $t-1$. Standard errors in parentheses are clustered at the commuting zone level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 7: Annual disbursement of ARRA funds.

	Total Funds		Green Funds	
Year	Award Amount (billion USD)	Share (%)	Award Amount (billion USD)	Share (%)
2009	194.80	70.93	31.89	63.12
2010	67.70	24.65	10.71	21.20
2011	11.80	4.30	7.90	15.64
2012	0.32	0.12	0.02	0.04

Notes: The table presents annual disbursements of ARRA funds and green ARRA funds. The data is from the NBER American Recovery and Reinvestment Act Federal Spending dataset. Green funds are mainly those allocated to the Department of Energy and the Environmental Protection Agency.

following model:

$$Y_{ict} = \alpha_0 \log \frac{\text{Green ARRA}_c}{\text{Population}_c} + \alpha_1 \text{Post}_t^{2009-2012} \times \log \frac{\text{Green ARRA}_c}{\text{Population}_c} + \alpha_2 \text{Post}_t^{2009-2012} + \alpha_3 \text{Post}_t^{2013} \times \log \frac{\text{Green ARRA}_c}{\text{Population}_c} + \alpha_4 \text{Post}_t^{2013} + \text{Year FE} + \text{City FE} + \varepsilon_{ict}, \tag{3}$$

where Y is the startup-level innovation outcomes for firm i founded in year t in commuting zone c (thus, each observation corresponds to a unique startup), $\text{Post}^{2009-2012}$ is an indicator for the initial four years following the passage of the ARRA (2009–2012), Post^{2013} is an indicator for the subsequent four years (2013–2017), and $\frac{\text{Green ARRA}}{\text{Population}}$ is the per capita green funding made available by the ARRA. To avoid issues related to truncation bias, we exclude startups founded after 2017, as that is the population for which intellectual property grants may have already been filed but yet to be rejected or assigned by the end of our sample. Online Appendix Table B.6 presents summary statistics of the Green ARRA funding amounts and the sample of startups.

Table 8 reports results on startups' green innovation activities. Our two coefficients of interest are those on our continuous treatment measures, $\text{Post}^{2009} \times \log \frac{\text{Green ARRA}}{\text{Population}}$. The $\text{Post}^{2009-2012}$-interacted measure tells us about the short-term effects of the influx of green ARRA funding, while the Post^{2013}-interacted measure tells us about the long-term effects of that funding. Panel (a) focuses on green trademarks results. In Column (1), the outcome variable is the ratio of green trademarks to total trademarks filed in the first two years of startup founding (i.e., year t to $t+1$). the estimated coefficient on $\text{Post}^{2009-2012} \times \log \frac{\text{Green ARRA}}{\text{Population}}$ in Column (1) is 0.004. An increase in log(Green ARRA/Population) from the 25th to the 75th percentile corresponds to a change of 3.13, implying an increase in the green trademark ratio of $0.004 \times 3.13 = 0.0125$. Relative to the unconditional mean of Y (0.033),[29] this corresponds to a $0.0125/0.033 = 37.9\%$ increase.[30]

In Columns (2)–(6) we calculate the ratio using the cumulative number of green trademarks and cumulative number of trademarks filed from founding through year n, with n ranging from 2 through 6. We see no evidence of differences between counties prior to the passage of the ARRA because the coefficient on $\log \frac{\text{Green ARRA}}{\text{Population}}$ is insignificant. We find statistically significant positive results during the four years in which the ARRA funding was disbursed, implying that the green ARRA funds led startups to engage in more early-stage green trademark filing. However, we do

29. Online Appendix Table B.8 presents summary statistics for the regression variables.

30. The economic interpretation of coefficients in an OLS regression is as follows. Consider the linear specification $Y = \alpha + \beta X + \varepsilon$. For an increase in X by one unit, the conditional expectation of Y increases by β units. More generally, if X increases by an interquartile range (IQR), i.e., from the 25th to the 75th percentile, the implied change in Y is $\Delta Y = \beta \times \text{IQR}$, and the implied percent change relative to a baseline level of Y is $(\beta \times \text{IQR})/Y$.

not find evidence of long run effects of the green ARRA on early-stage green trademark filing.

In Panel (b) we present results on startup patenting activity. The column meanings remain the same, but the outcome variable is now the proportion of early-stage patenting which is green. We find no evidence for short- or long-run effects of the green ARRA on startups' green patenting.

Table 8: Startup-level response to green ARRA funding.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Green Trademarks (ratio)						
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$	$t \to t+6$
$\text{Post}^{2009-2012} \times \log \frac{\text{Green ARRA}}{\text{Population}}$	0.004**	0.003**	0.003*	0.002*	0.002	0.003*
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Post 2013 $\times \log \frac{\text{Green ARRA}}{\text{Population}}$	-0.000	-0.001	-0.001	-0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$\log \frac{\text{Green ARRA}}{\text{Population}}$	0.003	0.002	0.002	0.003	0.003	0.003
	(0.004)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
Founding Year Fixed Effect	Y	Y	Y	Y	Y	Y
Headquarter Fixed Effect	Y	Y	Y	Y	Y	Y
Observations	12,598	19,937	24,643	28,107	30,813	32,841
R^2	0.106	0.084	0.077	0.071	0.066	0.068
Panel (b): Green Patents (ratio)						
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$	$t \to t+6$
$\text{Post}^{2009-2012} \times \log \frac{\text{Green ARRA}}{\text{Population}}$	-0.009	-0.005	-0.008	-0.004	-0.002	0.002
	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)
Post 2013 $\times \log \frac{\text{Green ARRA}}{\text{Population}}$	-0.002	0.000	0.001	0.005	0.006	0.007
	(0.008)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)
$\log \frac{\text{Green ARRA}}{\text{Population}}$	0.028*	0.001	-0.009	-0.002	0.001	0.004
	(0.015)	(0.014)	(0.013)	(0.010)	(0.009)	(0.009)
Founding Year Fixed Effect	Y	Y	Y	Y	Y	Y
Headquarter Fixed Effect	Y	Y	Y	Y	Y	Y
Observations	2,983	5,506	7,853	9,537	10,772	11,581
R^2	0.140	0.127	0.111	0.101	0.095	0.092

Notes: The table presents startup-level linear regression estimates of the effect of green ARRA funding on green innovation. In Panel (a) the outcome variable is the ratio of green trademarks. The specification is given in Equation (3). In Panel (b) the outcome variable is the ratio of green patents. In Column (1) the outcome variable is calculated over the first year of the startup. In Columns (2)–(6) the outcome variable is calculated over the first n years, with n ranging from 2 through 6. $\log \frac{\text{Green ARRA}}{\text{Population}}$ is the log of the dollar amount of green funding provided to the startup's commuting zone normalized by population in the year 2000. Post 2009–2012 is an indicator variable for the years immediately following the passage of the ARRA. Post 2013 is an indicator for years 2013 and onwards. Standard errors in parentheses are clustered at the commuting zone-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Next we examine the impact of the green ARRA on green startup formation in the treated

commuting zones. Using a Poisson regression, we estimate the following model:

$$\begin{aligned} Y_{ct} = \exp\big(&\alpha_0 \text{Post}_t^{2009-2012} \times \log \text{Green ARRA}_c + \alpha_1 \text{Post}_t^{2009-2012} \\ &+ \alpha_2 \text{Post}_t^{2013} \times \log \text{Green ARRA}_c + \alpha_3 \text{Post}_t^{2013} \\ &+ \alpha_4 \log \text{Green ARRA}_c + \text{Year FE} + \text{State FE} + \varepsilon_{ct}\big), \end{aligned} \tag{4}$$

where Y_{ct} is the number of green startups founded in commuting zone c in year t, $\text{Post}^{2009-2012}$ is an indicator for the initial four years following the passage of the ARRA (2009–2012), Post^{2013} is an indicator for the subsequent four years (2013–2017), and Green ARRA is the green funding made available by the ARRA to that commuting zone.[31] Given that our observed outcomes are commuting zone-aggregates, we use the overall amount of green funding spent in the commuting zone rather than a per-capita scaled amount of green funding.

In Table 9 we report results on commuting zone-level startup formation. Our two coefficients of interest are again those on our continuous treatment measures, $\text{Post}^{2009-2012} \times \log \text{Green ARRA}$ and $\text{Post}^{2013} \times \log \text{Green ARRA}$. As before, the interpretation of the $\text{Post}^{2009-2012}$-interacted regressor is that of the short-term effects of the green ARRA funding while the Post^{2013}-interacted regressor is that of the long-term effects of the funding. We find that, across several definitions of green startups, the green ARRA funding is associated with higher numbers of startups founded over the short-term in the counties that receive more funding.

In Panel (a), we report results on formation of the green-innovation startups. We show that, across several specifications, the coefficient on $\text{Post}^{2009-2012} \times \log \text{Green ARRA}$ remains positive. Next, we provide an economic interpretation of the coefficient of interest from our Poisson regression. In Column (2), the estimated coefficient on $\text{Post}^{2009-2012} \times \log \text{Green ARRA}$ is 0.094, and the interquartile-range change is 4.6. The implied percent change in green-innovation startups is therefore $\exp(0.094 \times 4.6) - 1 = 53.9\%$. Using the unconditional mean number of green-innovation startups (classified based on filings from t to $t+2$) of 0.99,[32] this corresponds to an increase of $0.99 \times (53.9\%)$ being approximately 0.53 additional green-innovation startups per commuting-zone-year. For the post-2013 period, the estimated coefficient in Column (2) is 0.158. Applying the same calculation yields $\exp(0.158 \times 4.6) - 1 = 106.4\%$, which corresponds to $0.99 \times (\exp(0.158 \times 4.6) - 1) \approx 1.05$ additional green-innovation startups per commuting-zone-year.

31. We assign commuting zones to states based on the largest state by population which overlaps with the commuting zone.

32. See Appendix Table B.7, which reports summary statistics for the Green ARRA regression variables

In Panel (b), we report results on the formation of self-labeled green startups. Column (1), the estimated coefficient on $\text{Post}^{2009-2012}$ is 0.07. Using an interquartile-range change of 5.2, this estimate implies a percent increase in self-labeled green startups of $\exp(0.07 \times 5.2) - 1 = 43.7\%$. Using the unconditional mean of N^{50} (9.3),[33] this translates into an increase from 9.3 to $9.3 \times \exp(0.07 \times 5.2) = 13.4$ green startups—that is, roughly 4.1 additional startups per commuting-zone-year.[34] This effect is economically meaningful: it equals 43.7% of the untreated baseline and, if persistent, implies about 20 additional green startups over a five-year horizon

Mirroring our results in Panel (a), we again find a significant and positive impact on long-term startup formation as well for the more heavily treated counties. Here, green startups are identified based on the intensity of green keywords in their business descriptions. We adopt increasingly strict cutoffs, with Column (1) classifying firms above the median intensity as green, Column (2) classifying firms above the 75^{th} percentile as green, and Column (3) classifying firms above the 90^{th} percentile as green. In Columns (4)–(6), we repeat this pattern of increasingly strict cutoffs, but use an industry-specific cutoff of green intensity rather than a common cutoff across all startups. for a treated commuting zone. For the post-2013 period, the estimated coefficient in Column (1) is 0.104. Applying the same calculation yields $\exp(0.104 \times 5.2) - 1 = 71.3\%$, corresponding to an increase from 9.3 to $9.3 \times \exp(0.104 \times 5.2) = 15.9$ green startups—about 6.7 additional startups per commuting-zone-year, or approximately 33 additional startups over five years.

5.3 Evidence from natural disasters and public attention to climate issues

After a major climate event, entrepreneurs and firm managers may update beliefs about the need for investment into climate adaptation and mitigation. This creates a plausible channel through which climate disasters can spur new green firm creation as well as green innovations. To study the quantitative importance of this channel, we examine startup formation and green innovations following natural disasters, defined as events in which a county suffers economic damages exceeding

33. Appendix Table B.8 presents summary statistics for the regression variables.

34. The economic interpretation of coefficients in a Poisson regression is as follows. Consider a Poisson model with a log link,

$$\ln \mathbb{E}[Y \mid X] = \alpha + \beta X.$$

For a commuting zone in which $\log(\textit{Green ARRA})$ increases from the 25th to the 75th percentile—i.e., by the interquartile range (IQR)—the conditional mean changes from $\exp(\alpha)$ to $\exp(\alpha + \beta\,\text{IQR})$. Therefore,

$$\frac{\mathbb{E}[Y \mid X + \text{IQR}]}{\mathbb{E}[Y \mid X]} = \exp(\beta\,\text{IQR}),$$

and the implied proportional change in Y is $\exp(\beta\,\text{IQR}) - 1$.

Table 9: Commuting zone-level response to green ARRA funding.

Panel (a): Green-innovation Startups (counts)						
	(1)	(2)	(3)	(4)	(5)	(6)
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$	$t \to t+6$
Post 2009–2012 × log Green ARRA	0.073	0.094***	0.067**	0.081***	0.072***	0.083***
	(0.046)	(0.031)	(0.026)	(0.024)	(0.023)	(0.024)
Post 2013 × log Green ARRA	0.157***	0.158***	0.141***	0.138***	0.129***	0.129***
	(0.043)	(0.034)	(0.029)	(0.026)	(0.025)	(0.024)
log Green ARRA	0.319***	0.303***	0.323***	0.325***	0.336***	0.332***
	(0.045)	(0.044)	(0.046)	(0.045)	(0.045)	(0.045)
State Fixed Effects	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	2,846	2,855	2,906	2,906	2,906	2,906
Pseudo-R^2	0.417	0.478	0.528	0.550	0.559	0.564
Panel (b): Self-labeled Green Startups (counts)						
	N^{50}	N^{75}	N^{90}	N^{50}_{ind}	N^{75}_{ind}	N^{90}_{ind}
Post 2009–2012 × log Green ARRA	0.070***	0.050***	0.042**	0.062***	0.055***	0.037**
	(0.017)	(0.016)	(0.018)	(0.018)	(0.017)	(0.017)
Post 2013 × log Green ARRA	0.104***	0.084***	0.066***	0.098***	0.081***	0.062***
	(0.020)	(0.019)	(0.019)	(0.020)	(0.019)	(0.019)
log Green ARRA	0.344***	0.339***	0.338***	0.365***	0.362***	0.364***
	(0.030)	(0.029)	(0.030)	(0.032)	(0.032)	(0.032)
State Fixed Effects	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	4,772	4,772	4,772	4,772	4,772	4,772
Pseudo-R^2	0.647	0.601	0.551	0.657	0.618	0.573

Notes: The table presents Poisson regression results for the impact of green ARRA funding on green startup formation. The specification is given in Equation (4). In Panel (a), the outcome variable is the count of new green innovation startups founded each year. In Column (1), the outcome variable classifies green innovation startups based on green patent or trademark filings in the first year post-founding. In Column (2) we expand the definition of green innovation startups to include firms that file a green patent or trademark in the first two years post-founding. In Columns (3)–(6) we further expand the definition to include the first n years, with n ranging from 3 to 6. In Panel (b), the outcome variable is counts of self-labeled green startups. In Column (1), the outcome variable is the count of new green startups founded each year, where green startups are those whose green keyword intensity is above the full-sample median. In Column (2), we use the 75^{th} percentile of green keyword intensity as the cutoff. In Column (3), we use the 90^{th} percentile of green keyword intensity as the cutoff. In Columns (4)–(6), we repeat the same sequence of cutoffs, but using industry-specific cutoffs rather than full sample cutoffs. Standard errors in parentheses are clustered at the commuting zone-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

one million dollars.

We use panel data on natural disasters in the continental U.S. to identify instances in which a county suffers economic damages exceeding one million dollars. To ensure that we are capturing the salience channel, we focus on startups in neighboring counties, rather than in the directly impacted county. This empirical strategy minimizes concerns about reverse-causation with regards to firm formation, in which we might expect impacted counties to have more firm entry due to financial losses by incumbent firms.

Our first set of results relates to startups' green innovations following a natural disaster. Using ordinary least squares regression, we estimate the following model:

$$Y_{ict} = \alpha_D \text{Disaster}^{\text{Adj}}_{c,t-1} + \alpha'_X X_{c,t-1} + \text{Year FE} + \text{County FE} + \varepsilon_{ict}, \tag{5}$$

where Y_{ict} is the startup-level innovation outcomes for startup i founded in year t in county c, $\text{Disaster}^{\text{Adj}}_{c,t-1}$ is an indicator for whether a county's neighboring counties c suffered a natural disaster in year $t-1$ exceeding 1 million USD in 2023 dollars, $X_{c,t-1}$ is a vector of county-year controls. We further include year and county fixed effects to absorb baseline disaster rates.

In Table 10 we report results on startups' green innovation activities. The coefficient of interest is the one on $\text{Disaster}^{\text{Adj}}_{c,t-1}$, which quantifies the impact of the natural disaster on subsequent startups' green innovations. Panel (a) presents results relating to green trademark filings. In Column (1), the outcome variable is the green trademark ratio — the proportion of green trademarks divided by all trademarks being filed from year t to $t+1$. In Columns (2)–(6) we extend the observation window to the 2^{nd} through 6^{th} year after the current year, where the numerator is the cumulative green trademark filings and the denominator is the cumulative trademark filings over the period (year t to $t+n$). There is a statistically significant and positive effect on green trademark filings by startups. Conditional on filing at least 1 trademark, startups that experience a county-adjacent natural disaster in the previous year on average increase the proportion of green trademarks by 0.6 percentage points in year t to $t+n$. This represents 23.1% of the unconditional sample average of 2.6 percent. This effect persists over the first five years of firm life. The estimated effects remain relatively stable at 0.3 percentage points if we expand the post-natural-disaster window from 2 to 5 years shown in Columns (2) through (5).

In Panel (b), we present analogous results where the outcome variable is now the ratio of green patent filings. Here, we find no evidence that the natural disaster lead to additional green

Table 10: Startup-level response to disasters.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Green Trademarks (ratio)						
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$	$t \to t+6$
$\text{Disaster}_{t-1}^{\text{Adj}}$	0.006*	0.003*	0.003**	0.002*	0.002**	0.001
	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
$\log \text{Population}_{t-1}$	-0.007	-0.063*	-0.077***	-0.056**	-0.033*	-0.025
	(0.036)	(0.032)	(0.026)	(0.022)	(0.019)	(0.016)
$\log \text{Income}_{t-1}$	-0.023	0.005	0.029	0.026	0.019	0.013
	(0.026)	(0.026)	(0.020)	(0.017)	(0.015)	(0.013)
Startup FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,436	24,140	34,084	41,829	48,054	53,224
Adj.-R^2	0.403	0.627	0.721	0.751	0.783	0.812
Panel (b): Green Patents (ratio)						
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$	$t \to t+6$
$\text{Disaster}_{t-1}^{\text{Adj}}$	-0.006	-0.009**	-0.005*	-0.002	-0.002	-0.002
	(0.006)	(0.004)	(0.003)	(0.002)	(0.002)	(0.002)
$\log \text{Population}_{t-1}$	-0.092	-0.089	-0.127	-0.059	-0.033	0.001
	(0.087)	(0.076)	(0.072)	(0.064)	(0.059)	(0.061)
$\log \text{Income}_{t-1}$	-0.023	0.036	0.016	-0.014	0.013	-0.005
	(0.061)	(0.055)	(0.056)	(0.051)	(0.045)	(0.046)
Startup FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,688	17,336	21,480	24,342	26,443	28,099
Adj.-R^2	0.582	0.715	0.765	0.799	0.823	0.835

Notes: The table presents OLS regression results on the startup-level impact of natural disasters on green innovation. The specification is given in Equation (5). In Panel (a), the outcome variable is the ratio of green trademarks to all trademarks filed in a given year. In Panel (b), the outcome variable is the ratio of green patents to all patents in a given year t. In Column (1), the ratio is calculated using all filings from year $(t \to t+1)$. In Columns (2)–(6), the ratio is calculated using all filings from the current year t to $t+n$ for n ranging from 2 through 6. $\text{Disaster}_{t-1}^{\text{Adj}}$ is an indicator variable for whether a county adjacent to the startup headquarters' county experienced a natural disaster in year $t-1$. County Population and Income, from the U.S. Bureau of Economic Analysis, are measured at the county level. Standard errors in parentheses are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

patenting activity, and even measure negative, albeit insignificant, impacts on green patenting. All specifications include county-year level controls for the county population and average income level, along with startup and year fixed effects.

Next, we examine the impact of a natural disaster on green startup formation in the treated (neighboring) counties. Using Poisson regression, we estimate the following model:

$$Y_{ct} = \exp\left(\alpha_D \text{Disaster}^{\text{Adj}}_{c,t-1} + \alpha'_X X_{c,t-1} + \text{Year FE} + \text{County FE} + \varepsilon_{ct}\right), \tag{6}$$

where Y_{ct} is now the number of green startups founded in county c in year t and other covariates are unchanged.

In Table 11 we report results on county-level startup formation following a natural disaster. In Panel (a) we first report results on the formation of green-innovation startups. Green-innovation startups are those filing a green patent or trademark within the first n years, where we allow n to vary from 1 to 6 years. In Column (1), counties in the current year adjacent to at least 1 county with damages exceeding 1 million USD in 2023 dollars caused by natural disasters in the previous year on average experience an increase in the creation of startups that have filed at least 1 green trademark or patent up to the founding years t and $t+1$ by 29.2%. Expanding this window since founding, we find similar magnitudes and statistical significance (all below the 5% level) in Columns (2)–(6), which corresponds to 1 to 5 years since founding.

In Panel (b) of Table 11, we report results on self-labeled green startups. We find mixed statistical evidence for increased green startup formation over several definitions of self-labeled green startups. In Column (1) we identify green startups as those whose business descriptions feature an intensity of environmental keywords above the median of firms in our sample. We observe a positive and statistically significant effect. Counties that are exposed to at least 1 county-adjacent natural disaster experience an average increase of $100 \times (\exp(0.088) - 1) = 9.2\%$ in newly founded self-labeled green startups, as defined has having green keyword intensity in their business description above the sample median, in the years t and $t+1$. This is also the case in Column (2) where we use a stricter 75^{th} percentile cutoff rather than the median. In Column (3) we use the 90^{th} percentile and do not find a statistically significant effect. In Columns (4)–(6) we use industry-specific cutoff thresholds, and find consistently positive effects of natural disasters on green startup formation in neighboring counties. In Columns (5) and (6) where we define self-labeled green startups as those with green keyword intensity in their business descriptions above the 75 and 90

percentile in their respective industries, we find that counties exposed to at least 1 county-adjacent natural disaster increase the average number of self-labeled green startups founded by 12% and 10%, respectively. These effects are statistically significant at the 5% level. All specifications include county-year level controls for the county population and average income level, along with county and year fixed effects.

5.4 Summary of factors for green innovations and entrepreneurship

In this section, we analyze three potential factors that may explain the surge of green entrepreneurship all over the U.S.: education, government policies, and public attention to climate issues. To examine the role of education, we first collect an extensive list of green research institutes in U.S. universities and their establishment years. We find that once such institutes were established, local new startups' proportion of green trademarks increases by 0.8 percentage points (which corresponds to 21% of sample mean), and the number of local startups with green innovations grows by 25.4%. All of these statistics confirm the role of education in the pursuit of sustainable growth. We next consider governments' green investment from the American Recovery and Reinvestment Act (ARRA) between 2009 and 2012. Startups located in areas receiving more green investment per capita file a significantly higher proportion of green trademarks, with an increase of 1.5% percentage points. Moreover, the number of new startups with green innovations grows by 53.9% in the funding period. Our evidence thus supports the effectiveness of such funding from the perspective of local entrepreneurship.

Furthermore, we use the incidences of natural disasters that occurred in adjacent counties to capture the local's attention to climate issues. Our empirical results suggest that local new startups' proportion of green trademarks increases by 0.6 percentage points (corresponding to 23.1% of sample mean) right after such an incidence. Moreover, the number of startups with green innovations grows by 29.5% right after such an incidence. All these results point to the importance of public attention in shaping the landscape of green entrepreneurship.

Table 11: County-level response to disasters.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Green-innovation Startups (counts)						
	$t \to t+1$	$t \to t+2$	$t \to t+3$	$t \to t+4$	$t \to t+5$	$t \to t+6$
$\text{Disaster}_{t-1}^{\text{Adj}}$	0.256**	0.262**	0.351***	0.265***	0.299***	0.305***
	(0.116)	(0.113)	(0.104)	(0.088)	(0.088)	(0.092)
$\log \text{Population}_{t-1}$	-4.401***	-3.968***	-3.873***	-3.498***	-3.180***	-2.909***
	(1.592)	(1.403)	(1.251)	(1.137)	(1.078)	(1.042)
$\log \text{Income}_{t-1}$	2.306*	2.137**	2.155**	2.140***	2.125***	2.110***
	(1.346)	(1.009)	(0.856)	(0.753)	(0.743)	(0.706)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,018	1,348	1,680	1,886	2,088	2,310
Pseudo-R^2	0.439	0.528	0.607	0.637	0.654	0.662
Panel (b): Self-labeled Green Startups (counts)						
	N^{50}	N^{75}	N^{90}	N_{ind}^{50}	N_{ind}^{75}	N_{ind}^{90}
$\text{Disaster}_{t-1}^{\text{Adj}}$	0.088*	0.100*	0.048	0.083	0.115**	0.096**
	(0.047)	(0.060)	(0.069)	(0.068)	(0.054)	(0.049)
$\log \text{Population}_{t-1}$	-2.530***	-2.431***	-2.358***	-2.540***	-2.367***	-2.264***
	(0.682)	(0.649)	(0.578)	(0.667)	(0.678)	(0.709)
$\log \text{Income}_{t-1}$	1.841***	1.685***	1.473***	1.633***	1.697***	1.662***
	(0.552)	(0.522)	(0.498)	(0.525)	(0.537)	(0.558)
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,097	9,097	5,653	5,024	7,890	11,099
Pseudo-R^2	0.790	0.700	0.576	0.625	0.734	0.809

Notes: The table presents Poisson regression results on the county-level impact of natural disasters on green startup formation. The specification is given in Equation (6). In Panel (a) the outcome variable is the count of green innovation startups. In Column (1) we classify any startup that files a green patent or trademark in the first year of its life as a green innovation startup. In Column (2) we expand the definition to any startup that files a green patent or trademark in the first two years of its life. In Columns (3)–(6) we expand the definition to any startup that files a green patent or trademark in the first n years, with n ranging from 3 to 6. In Panel (b) the outcome variable is the count of self-labeled green startups. In Column (1), the outcome variable is the count of new green startups founded each year, where green startups are those whose green keyword intensity is above the full-sample median. In Column (2), we use the 75^{th} percentile of green keyword intensity as the cutoff. In Column (3), we use the 90^{th} percentile of green keyword intensity as the cutoff. In Columns (4)–(6), we repeat the same sequence of cutoffs, but using industry-specific cutoffs rather than full sample cutoffs. $\text{Disaster}_{t-1}^{\text{Adj}}$ is an indicator variable for whether a county adjacent to the startup headquarters' county experienced a natural disaster in year $t-1$. Population and Income are measured at the county level. Standard errors in parentheses are clustered at the county level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

6 Conclusion

A large-scale investigation of U.S. startups' environmental engagement has been a void in the literature. In this paper, we take the challenge and construct a comprehensive dataset of green entrepreneurship in the U.S. by linking Pitchbook to Revelio Labs data for startups' and founders' profiles and to USPTO's patent and trademark data for those startups' innovations. We then use the classifications in the latest research to identify patents and trademarks specific to environment and climate issues, which constitute our green innovations. This dataset enables us to analyze the development of green innovations created by all startups and the time-series and geographic distribution of green startups.

We first document strong growth in all startups' green innovation activity over time. The percentage of green patent filings increased from 15% in 1980 to 25% in 2023 and the percentage of green trademark filings steadily grew from 2% in 1980 to 5% in 2023. In terms of geographic distribution, more green startups are formed in states with larger populations (e.g., California, New York, Texas, Massachusetts, and Pennsylvania). More interestingly, the proportion of startups with green innovations also increases in these states in time, suggesting that the rise of green entrepreneurship is a general pattern across all states.

We further consider startups' intensity of environment-related terms in their business descriptions in Pitchbook, and use that information to define self-labeled green startups. We find that these startups indeed create more green patents and trademarks. These creations are more likely to occur earlier in their life cycles. Our evidence thus suggests that these self-labeled green startups deliver on their business promises.

Analyzing three potential factors driving green startup innovation and formation, we find that the local entrepreneurial ecosystem, funding availability, and local natural disasters all drive local startups' green innovation. We proxy the "green-ness" of the local entrepreneurial ecosystem with the opening of a green research school, institute or center, as universities are an agglomeration hub for entrepreneurs, researchers and investors. To analyze the impact of federal funding, we study the impact of the green funds made available in the ARRA. Finally, we study the local response to significant natural disasters. Each factor is associated with a greater mix of green innovations from local startups along with higher rates of green startup formation.

Our use of startups' green patents and trademarks demonstrates new research opportunities for researchers to take broader and more quantitative approaches to analyze startups' activity

in the sustainability dimension. In addition, the more pronounced engagement of startups in green trademarking suggests that, because they prioritize commercialization, their efforts are more inclined to focus on launching new product lines rather than developing new technologies.

Moreover, our comprehensive analysis of the dynamics and explanatory factors of green entrepreneurship offers rich policy implications. First, given the resource constraints of startups, intellectual property policies that protect their green innovations can be a factor in sustainable growth.[35] Second, although they are relatively small in size, startups' growing involvement in green innovation signals a major forthcoming shift in the business landscape. Such momentum and regime-shift in entrepreneurship highlight the necessity of synchronizing governments' policies on SMEs and environmental protection. Finally, the surge of green entrepreneurship can be attributed to multifaceted driving forces, which confirms the complementarity of different policy tools from the SMEs' perspective and calls for further in-depth investigation into the optimization of policy portfolios.

35. An example is the USPTO's Climate Change Mitigation Pilot Program that offers faster review and approval process for patent applications on climate-related innovations: https://www.uspto.gov/patents/laws/patent-related-notices/climate-change-mitigation-pilot-program.

References

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. 2012. The environment and directed technical change. *American Economic Review* 102, no. 1 (February): 131–166. (Cited on page 6).

Acemoglu, Daron, Ufuk Akcigit, Douglas Hanley, and William Kerr. 2016. Transition to clean technology. *Journal of Political Economy* 124, no. 1 (February): 52–104. (Cited on page 6).

Aghion, Philippe, Antoine Dechezleprêtre, David Hémous, Ralf Martin, and John Van Reenen. 2016. Carbon taxes, path dependency, and directed technical change: evidence from the auto industry. *Journal of Political Economy* 124, no. 1 (February): 1–51. (Cited on page 6).

Akey, Pat, and Ian Appel. 2020. The limits of limited liability: evidence from industrial pollution. *The Journal of Finance* 76, no. 1 (October): 5–55. (Cited on page 5).

Aldy, Joseph E. 2013. Policy monitora preliminary assessment of the american recovery and reinvestment act's clean energy package. *Review of Environmental Economics and Policy* 7, no. 1 (January): 136–155. (Cited on pages 3, 33).

Azoulay, Pierre, Joshua S. Graff Zivin, and Gustavo Manso. 2011. Incentives and creativity: evidence from the academic life sciences. *The RAND Journal of Economics* 42, no. 3 (September): 527–554. (Cited on page 6).

Azoulay, Pierre, Misty Heggeness, and Jennifer Kao. 2025. Medical research and health care finance: evidence from academic medical centers. *Research Policy* 54, no. 2 (March): 105163. (Cited on page 6).

Babina, Tania, Alex Xi He, Sabrina T Howell, Elisabeth Ruth Perlman, and Joseph Staudt. 2023. Cutting the innovation engine: how federal funding shocks affect university patenting, entrepreneurship, and publications. *The Quarterly Journal of Economics* 138, no. 2 (January): 895–954. (Cited on page 6).

Baier, Philipp, Marc Berninger, and Florian Kiesel. 2020. Environmental, social and governance reporting in annual reports: a textual analysis. *Financial Markets, Institutions & Instruments* 29 (3): 93–118. (Cited on pages 2, 9).

Bellon, Aymeric, J Anthony Cookson, Erik P Gilje, and Rawley Z Heimer. 2021. Personal wealth, self-employment, and business ownership. Edited by Stijn Van Nieuwerburgh. *The Review of Financial Studies* 34, no. 8 (April): 3935–3975. (Cited on pages 1, 5).

Bénabou, Roland, and Jean Tirole. 2009. Individual and corporate social responsibility. *Economica* 77, no. 305 (December): 1–19. (Cited on page 1).

Bernstein, Shai, Richard R Townsend, and Ting Xu. 2023. Flight to safety: how economic downturns affect talent flows to startups. Edited by Holger Mueller. *The Review of Financial Studies* 37, no. 3 (September): 837–881. (Cited on page 1).

Bhattacharya, C.B., and Sankar Sen. 2003. Consumer–company identification: a framework for understanding consumers' relationships with companies. *Journal of Marketing* 67, no. 2 (April): 76–88. (Cited on page 1).

Bolton, Patrick, Marcin T. Kacperczyk, and Moritz Wiedemann. 2022. The co2 question: technical progress and the climate crisis. *SSRN Electronic Journal,* (cited on pages 2, 11).

Brander, James A., Qianqian Du, and Thomas Hellmann. 2014. The effects of government-sponsored venture capital: international evidence. *Review of Finance* 19, no. 2 (March): 571–618. (Cited on page 6).

Brown, James R, Gustav Martinsson, and Christian Thomann. 2022. Can environmental policy encourage technical change? emissions taxes and r&d investment in polluting firms. Edited by David Denis. *The Review of Financial Studies* 35, no. 10 (January): 4518–4560. (Cited on page 5).

Chen, Yi-Chun, Mingyi Hung, and Yongxiang Wang. 2018. The effect of mandatory csr disclosure on firm profitability and social externalities: evidence from china. *Journal of Accounting and Economics* 65, no. 1 (February): 169–190. (Cited on page 6).

Chiu, Wan-Chien, Po-Hsuan Hsu, (Joy) Tianjiao Tong, and Kai Li. 2024. Green products *, (cited on pages 4, 5, 11).

Cohen, Lauren H., Umit G. Gurun, and Quoc H. Nguyen. 2026. The esg-innovation disconnect: evidence from green patenting. *Management Science* (January). (Cited on page 5).

Darendeli, Alper, Kelvin K. F. Law, and Michael Shen. 2022. Green new hiring. *Review of Accounting Studies* 27, no. 3 (August): 986–1037. (Cited on page 5).

Doblinger, Claudia, Kavita Surana, and Laura Diaz Anadon. 2019. Governments as partners: the role of alliances in u.s. cleantech startup innovation. *Research Policy* 48, no. 6 (July): 1458–1475. (Cited on page 6).

Downar, Benedikt, Jürgen Ernstberger, Stefan Reichelstein, Sebastian Schwenen, and Aleksandar Zaklan. 2021. The impact of carbon disclosure mandates on emissions and financial operating performance. *Review of Accounting Studies* 26, no. 3 (August): 1137–1175. (Cited on page 6).

Duchin, Ran, Janet Gao, and Qiping Xu. 2024. Sustainability or greenwashing: evidence from the asset market for industrial pollution. *The Journal of Finance* 80, no. 2 (December): 699–754. (Cited on page 5).

Engelberg, Joey, Jorge Guzman, Runjing Lu, and William Mullins. n.d. Partisan entrepreneurship. Forthcoming. *Journal of Finance,* (cited on page 5).

European Union Intellectual Property Office, and. 2021. *Green eu trade marks – analysis of goods and services specifications, 1996-2020.* European Union Intellectual Property Office. (Cited on page 11).

Ewens, Michael, and Matt Marx. 2023. Firm age and invention: an open access dataset. Working Paper. (Cited on page 5).

Feldman, Maryann P., and Maryellen R. Kelley. 2003. Leveraging research and development: assessing the impact of the u.s. advanced technology program. *Small Business Economics* 20, no. 2 (March): 153–165. (Cited on page 6).

Flammer, Caroline. 2021. Corporate green bonds. *Journal of Financial Economics* 142, no. 2 (November): 499–516. (Cited on page 5).

Gans, Joshua, and Scott Stern. 2003. When does funding research by smaller firms bear fruit?: evidence from the sbir program*. *Economics of Innovation and New Technology* 12, no. 4 (August): 361–384. (Cited on page 6).

Gonzalez-Uribe, Juanita, and Michael Leatherbee. 2017. The effects of business accelerators on venture performance: evidence from start-up chile. *The Review of Financial Studies* 31, no. 4 (September): 1566–1603. (Cited on page 6).

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. 2022. Can technology solve the principal-agent problem? evidence from china's war on air pollution. *American Economic Review: Insights* 4, no. 1 (March): 54–70. (Cited on page 6).

Guzman, Jorge, Jean Joohyun Oh, and Ananya Sen. 2020. What motivates innovative entrepreneurs? evidence from a global field experiment. *Management Science* 66, no. 10 (October): 4808–4819. (Cited on page 5).

Haščič, Ivan, and Mauro Migotto. 2015. Measuring environmental innovation using patent data. *OECD Environment Working Papers* (June).

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar. 2020. Can unemployment insurance spur entrepreneurial activity? evidence from france. *The Journal of Finance* 75, no. 3 (February): 1247–1285. (Cited on page 5).

Horbach, Jens. 2008. Determinants of environmental innovation–new evidence from german panel data sources. *Research Policy* 37, no. 1 (February): 163–173. (Cited on page 5).

Howell, Sabrina T. 2017. Financing innovation: evidence from r&d grants. *American Economic Review* 107, no. 4 (April): 1136–1164. (Cited on page 6).

———. 2024. Government intervention in innovation. *Annual Review of Financial Economics* 16, no. 1 (November): 367–390. (Cited on page 6).

Hsu, Po-Hsuan, Kai Li, and Yihui Pan. 2025. The eco-gender gap in boardrooms. *Management Science* 0 (0). (Cited on page 5).

Hsu, Po-Hsuan, Kai Li, and Chi-Yang Tsou. 2023. The pollution premium. *The Journal of Finance* 78, no. 3 (April): 1343–1392. (Cited on page 5).

Hurst, Erik, and Annamaria Lusardi. 2004. Liquidity constraints, household wealth, and entrepreneurship. *Journal of Political Economy* 112, no. 2 (April): 319–347. (Cited on page 5).

Hurst, Erik, and Benjamin Wild Pugsley. 2011. *What do small businesses do?* May. (Cited on page 1).

Ivanov, Ivan T, Mathias S Kruttli, and Sumudu W Watugala. 2023. Banking on carbon: corporate lending and cap-and-trade policy. Edited by Ralph Koijen. *The Review of Financial Studies* 37, no. 5 (December): 1640–1684. (Cited on page 5).

Kacperczyk, Marcin T., and Jose-Luis Peydro. 2021. Carbon emissions and the bank-lending channel. *SSRN Electronic Journal,* (cited on page 5).

Klausmann, Johannes, Philipp Krueger, and Pedro Matos. 2024. The green transition: evidence from corporate green revenues, (cited on page 5).

Krueger, Philipp, Zacharias Sautner, and Laura T Starks. 2020. The importance of climate risks for institutional investors. Edited by Andrew Karolyi. *The Review of Financial Studies* 33, no. 3 (February): 1067–1111. (Cited on page 5).

Krüger, Philipp. 2015. Corporate goodness and shareholder wealth. *Journal of Financial Economics* 115, no. 2 (February): 304–329. (Cited on page 5).

Lanzi, Elisa, Elena Verdolini, and Ivan Haščič. 2011. Efficiency-improving fossil fuel technologies for electricity generation: data selection and trends. *Energy Policy* 39, no. 11 (November): 7000–7014.

Lerner, Josh. 2002. When bureaucrats meet entrepreneurs: the design of effective 'public venture capital' programmes. *The Economic Journal* 112, no. 477 (February): F73–F84. (Cited on page 6).

———. 2009. *Boulevard of broken dreams: why public efforts to boost entrepreneurship and venture capital have failed–and what to do about it.* Princeton University Press. (Cited on page 6).

Levine, Ross, and Yona Rubinstein. 2017. Smart and illicit: who becomes an entrepreneur and do they earn more? *The Quarterly Journal of Economics* 132 (2): pp. 963–1018. (Cited on page 1).

Lin, Yan, Rui Shen, Jasmine Wang, and Y Julia Yu. 2024. Global evolution of environmental and social disclosure in annual reports. *Journal of Accounting Research* 62 (5): 1941–1988. (Cited on pages 2, 9).

Meier, Jean-Marie, Henri Servaes, Jiaying Wei, and Steven Chong Xiao. 2022. Do consumers care about esg? evidence from barcode-level sales. *SSRN Electronic Journal,* (cited on page 5).

Miyamoto, Mai, and Kenji Takeuchi. 2019. Climate agreement and technology diffusion: impact of the kyoto protocol on international patent applications for renewable energy technologies. *Energy Policy* 129 (June): 1331–1338. (Cited on page 15).

Newell, R. G., A. B. Jaffe, and R. N. Stavins. 1999. The induced innovation hypothesis and energy-saving technological change. *The Quarterly Journal of Economics* 114, no. 3 (August): 941–975. (Cited on page 6).

Pencle, Nadra, and Irina Mălăescu. 2016. What's in the words? development and validation of a multidimensional dictionary for csr and application using prospectuses. *Journal of Emerging Technologies in Accounting* 13 (2): 109–127. (Cited on pages 2, 9).

Popp, David. 2002. Induced innovation and energy prices. *American Economic Review* 92, no. 1 (February): 160–180. (Cited on page 6).

———. 2003. Pollution control innovations and the clean air act of 1990. *Journal of Policy Analysis and Management* 22, no. 4 (September): 641–660. (Cited on page 6).

Popp, David, Richard G. Newell, and Adam B. Jaffe. 2010. Energy, the environment, and technological change. In *Handbook of the economics of innovation, volume 2,* 873–937. Elsevier. (Cited on page 6).

Popp, David, Francesco Vona, Giovanni Marin, and Ziqiao Chen. 2021. The employment impact of a green fiscal push: evidence from the american recovery and reinvestment act. *Brookings Papers on Economic Activity,* 1–49. (Cited on pages 3, 12, 33).

Preston, Anne E. 1989. The nonprofit worker in a for-profit world. *Journal of Labor Economics* 7, no. 4 (October): 438–463. (Cited on page 1).

Sauermann, Henry. 2017. Fire in the belly? employee motives and innovative performance in start-ups versus established firms. *Strategic Entrepreneurship Journal* 12, no. 4 (October): 423–454. (Cited on page 1).

Sautner, Zacharias, Laurence van Lent, Grigory Vilkov, and Ruishen Zhang. 2023a. Firm-level climate change exposure. *The Journal of Finance* 78, no. 3 (March): 1449–1498. (Cited on page 5).

———. 2023b. Pricing climate change exposure. *Management Science* 69, no. 12 (December): 7540–7561. (Cited on page 5).

Schumpeter, Joseph A. 1934. *The theory of economic development.* Harvard University Press. (Cited on page 1).

Sen, Sankar, and C.B. Bhattacharya. 2001. Does doing good always lead to doing better? consumer reactions to corporate social responsibility. *Journal of Marketing Research* 38, no. 2 (May): 225–243. (Cited on page 1).

Starks, Laura T., Parth Venkat, and Qifei Zhu. 2026. Corporate esg profiles and investor horizons. *The Journal of Finance* (January). (Cited on page 5).

Sun, Yuan, Xuan Tian, and Yuanchen Yang. 2024. A robust green patent database: a new dataset of green patents through large language models. In *Innovation information initiative technical working group meeting.* Organized by Matt Marx; supported by the Alfred P. Sloan Foundation. Hybrid meeting (in-person and Zoom). Royal Sonesta Hotel, Longfellow Room. Cambridge, MA, December 6–7, 2024. (Cited on pages 2, 11).

Tomar, Sorabh. 2023. Greenhouse gas disclosure and emissions benchmarking. *Journal of Accounting Research* 61, no. 2 (March): 451–492. (Cited on page 6).

Tonin, Mirco, and Michael Vlassopoulos. 2010. Disentangling the sources of pro-socially motivated effort: a field experiment. *Journal of Public Economics* 94, nos. 11–12 (December): 1086–1092. (Cited on page 1).

Toole, Andrew A., and Calum Turvey. 2007. How does initial public financing influence private incentives for follow-on investment in early-stage technologies? *The Journal of Technology Transfer* 34, no. 1 (December): 43–58. (Cited on page 6).

Xu, Qiping, and Taehyun Kim. 2021. Financial constraints and corporate environmental policies. Edited by Wei Jiang. *The Review of Financial Studies* 35, no. 2 (May): 576–635. (Cited on page 5).

Online Appendix

A Data Discussion

A.1 Constructing the Green Trademark dataset

We download the case file for all registered trademarks from the USPTO,[36] including a mark's identification, mark words, application date, registration date, Nice Class(es), owner name, and owner address. Our sample contains USPTO marks registered by corporations (legal entities, not individuals) between 1980 and 2023.

A.1.1 The EUIPO green trademark taxonomy

In 2019, the European Commission established action on climate change as a priority. In 2021, the European Union Intellectual Property Office (EUIPO) (2021) released its first version of the green trademark taxonomy, reflecting the increasing frequency with which the identification of goods and services of EU trademarks is related to environmental protection and sustainability.

The EUIPO begins by searching its Harmonized Database (HDB), introduced in 1996, for terms related to environmental protection and sustainability. Once this Harmonized Green Terms inventory is established, the EUIPO employs an algorithm that combines machine learning with expert intervention to generate 375 green expressions for identifying green marks. The full EUIPO green trademark taxonomy is provided in Annex 2 of the same study (European Union Intellectual Property Office 2021).

Each of those green expressions is assigned a reference number ("Ref.") ranging from 1 to 375. To adapt the EUIPO green expressions to the U.S. context, we modify them using the British-American-English conversion.[37] In our study, we call this green term classifier the "EUIPO green trademark taxonomy."

The EUIPO green trademark taxonomy assigns green terms into 9 different groups: (1) Agriculture, (2) Climate Change, (3) Energy Conservation, (4) Energy Production, (5) Environmental Awareness, (6) Pollution Control, (7) Reusable, (8) Transportation, and (9) Waste Management. Table A.1 contains five examples from 375 green expressions and their corresponding assigned groups shown in the last column.

A.1.2 Identifying green marks from the USPTO

The EUIPO green trademark taxonomy is developed to identify green trademarks as long as a mark's goods and/or services identification is in English. To account for our context of studying green marks applied in the U.S., we modify the taxonomy and describe our detailed procedure below. We provide examples in Section A.1.3.

Step 1 We collect detailed trademark data from the USPTO. For our purpose, we extract G&S identification and parse that piece of textual information. Specifically, the textual description for each unique Nice Class is separated by semicolons, which gives us the basic unit and is called a

36. Available at https://www.uspto.gov/ip-policy/economic-research/research-datasets/trademark-case-files-dataset.

37. Some examples of our modifications are as follows: "fertilisers" becomes "fertilizers" (Ref. 209); "refuelling" becomes "refueling" (Ref. 183); "demineralising" becomes "demineralizing" (Ref. 282); "deodorising" becomes "deodorizing" (Ref. 6 and Ref. 107).

Ref.	Green expression	Group
11	+bacteria +water.treatment	Pollution Control
120	+fuel +ethanol −pharmaceutical −topical −1	Energy Production
183	+rechargeable −acidulated −telephone −computer −fire.extinguisher −refuelling −sweepers	Energy Conservation
140	+hydrogen +vehicle	Transportation
202	+reusable +bags	Reusable

Table A.1: Some examples of green expressions. Row 1 corresponds to Ref. 11 and means that a term is green if it contains the word 'bacteria' and the phrase 'water treatment' (together and in this order). Ref. 120 means that a term is green if it contains the word 'fuel' and the word 'ethanol' and does not contain the word 'pharmaceutical' or the word 'topical', *conditional* on that mark not registering in Nice Class 1 Chemicals. Ref. 183 means that a term is green if it contains the word 'rechargeable' and does not contain the word 'acidulated' or the word 'telephone' or the word 'computer' or the phrase 'fire extinguisher' (together and in this order) or the word 'refuelling' or the word 'sweepers'. Ref. 140 means that a term is green if it contains the word 'hydrogen' and the word 'vehicle'. Ref. 202 means that a term is green if it contains the word 'reusable' and the word 'bags'.

'term'.[38] A 'term' is 'the atomic definition of a product' in the G&S identification that contains a Nice Class and words or phrases to describe the product (service) (European Union Intellectual Property Office 2021, p. 12), and will be the basic data unit for our analysis.

Step 2 We standardize the G&S description at term-level by removing stop words, punctuations, and stemming.[39] We tokenize the textual information.[40]

Step 3 We perform the same preprocessing step for the 375 green expressions in the EUIPO green trademark taxonomy.[41]

A.1.3 Examples of green marks identified from the USPTO

In this section, we demonstrate several examples from four companies: Streamline Innovations Inc., EnergyHub, Inc., RECYCLEBANK LLC, and Xpansiv Data Systems, Inc.

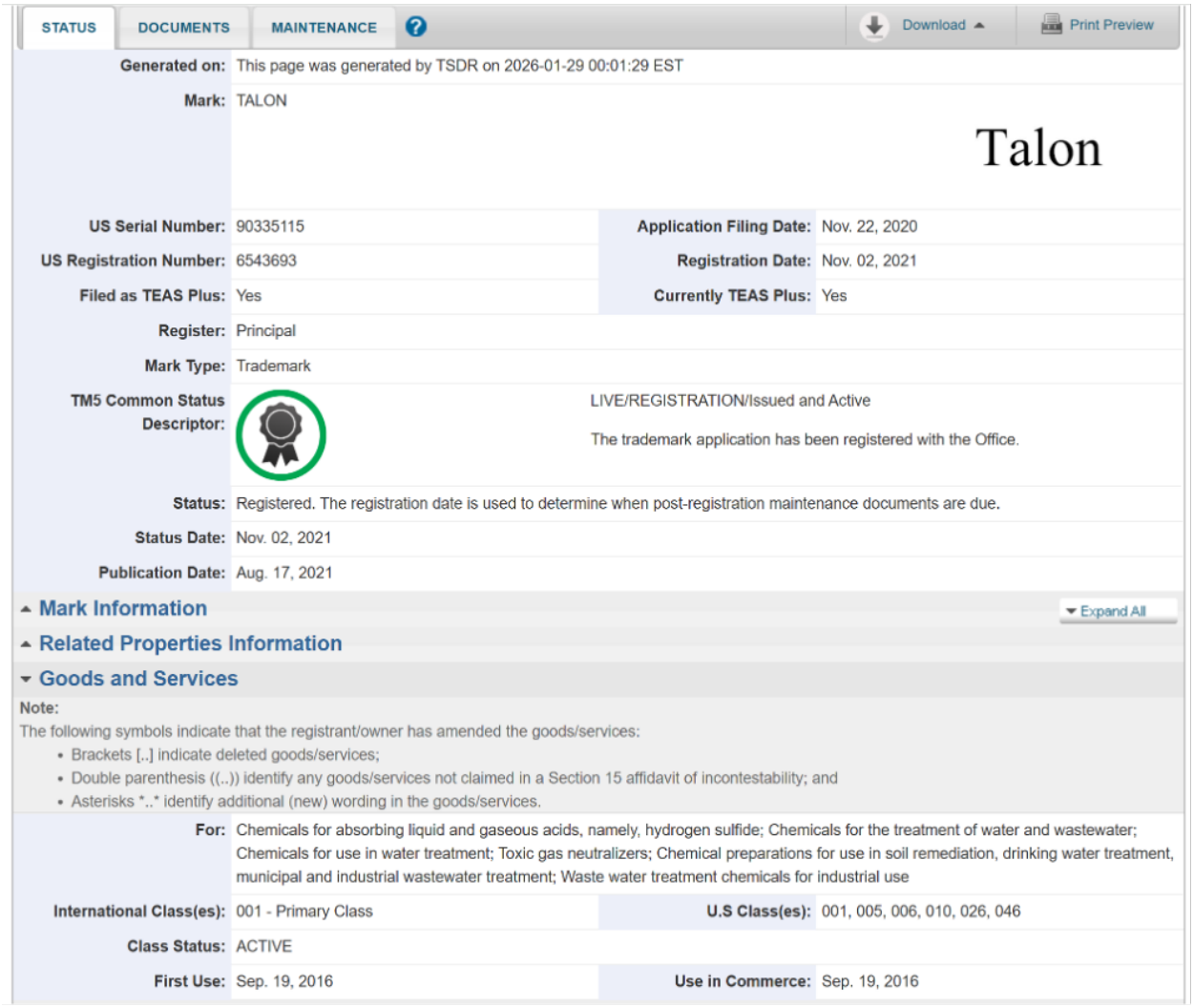
38. A trademark can apply under multiple Nice Classes.

39. We use Python and its `stop-words` and `pystemmer` packages. Stop words are a set of commonly used words in a language. Examples of stop words in English are 'a', 'the', 'is', and 'are'.

40. After tokenization, single words are separated. For compound words, we implement tokenization in two ways. The first approach is to remove hyphen; for example, "self-balancing" becomes "selfbalancing", and "eco-friendly" becomes "ecofriendly". The second approach is to replace the hyphen with a space; for example, "electric-motor" becomes "electric" and "motor". All compound words are tokenized using both approaches.

41. We perform the same preprocessing step—removing stop words and punctuations and stemming—to terms to be classified as well as the 375 green expressions in the EUIPO green trademark taxonomy.

Company 1: Streamline Innovations Inc. Streamline Innovations Inc. applied for this trademark on November 22, 2020, and registered it on November 2, 2021. The Word Mark is "TALON" with the registration number 6543693.



STATUS | DOCUMENTS | MAINTENANCE | Download | Print Preview

Generated on:	This page was generated by TSDR on 2026-01-29 00:01:29 EST
Mark:	TALON
	Talon
US Serial Number:	90335115
Application Filing Date:	Nov. 22, 2020
US Registration Number:	6543693
Registration Date:	Nov. 02, 2021
Filed as TEAS Plus:	Yes
Currently TEAS Plus:	Yes
Register:	Principal
Mark Type:	Trademark
TM5 Common Status Descriptor:	LIVE/REGISTRATION/Issued and Active The trademark application has been registered with the Office.
Status:	Registered. The registration date is used to determine when post-registration maintenance documents are due.
Status Date:	Nov. 02, 2021
Publication Date:	Aug. 17, 2021

▲ Mark Information — Expand All

▲ Related Properties Information

▼ Goods and Services

Note:
The following symbols indicate that the registrant/owner has amended the goods/services:
- Brackets [..] indicate deleted goods/services;
- Double parenthesis ((..)) identify any goods/services not claimed in a Section 15 affidavit of incontestability; and
- Asterisks *..* identify additional (new) wording in the goods/services.

For:	Chemicals for absorbing liquid and gaseous acids, namely, hydrogen sulfide; Chemicals for the treatment of water and wastewater; Chemicals for use in water treatment; Toxic gas neutralizers; Chemical preparations for use in soil remediation, drinking water treatment, municipal and industrial wastewater treatment; Waste water treatment chemicals for industrial use
International Class(es):	001 - Primary Class
U.S Class(es):	001, 005, 006, 010, 026, 046
Class Status:	ACTIVE
First Use:	Sep. 19, 2016
Use in Commerce:	Sep. 19, 2016

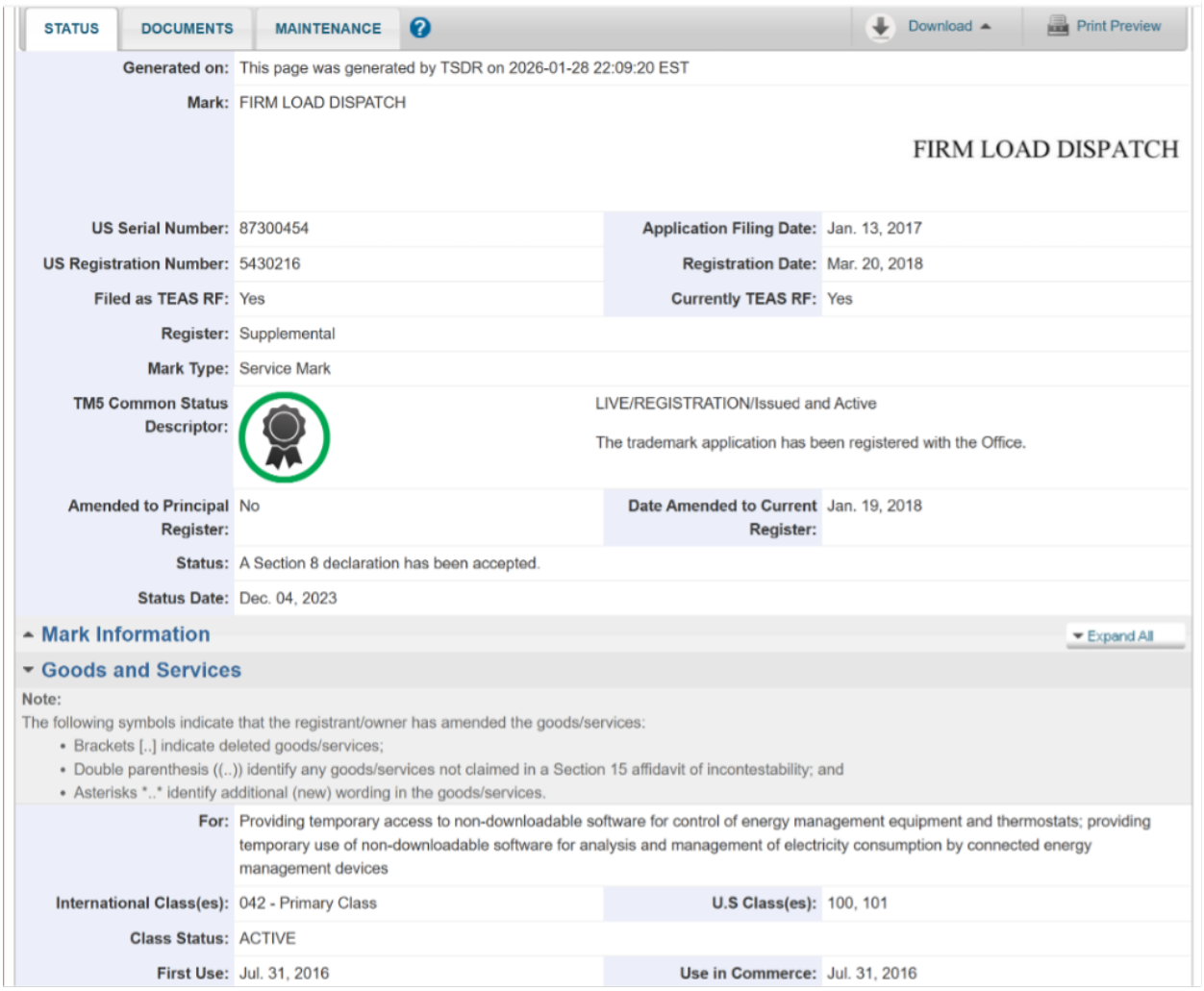
In this example, six terms in the goods and services identification are extracted and shown below. We match each term with any of the 375 green expressions. Five terms are matched to different green expressions shown as follows.

Registration number (Registration date)	Nice Class	Green Expression Ref.	Mark Word	G&S identification at the term level
3335977 (20-Mar-2018)	1	Not matched	TALON	Chemicals for absorbing liquid and gaseous acids, namely, hydrogen sulfide;
3335977 (20-Mar-2018)	1	254	TALON	Chemicals for the treatment of water and wastewater;
3335977 (20-Mar-2018)	1	282	TALON	Chemicals for use in water treatment;
3335977 (20-Mar-2018)	1	240	TALON	Toxic gas neutralizers;
3335977 (20-Mar-2018)	1	275	TALON	Chemical preparations for use in soil remediation, drinking water treatment, municipal and industrial wastewater treatment;
3335977 (20-Mar-2018)	1	273	TALON	Waste water treatment chemicals for industrial use;

These green terms are matched to Ref. 254, 282, 240, 275, and 273 shown below,

Group	Category	Ref.	Green expression
Pollution Control	Water purification	254	+treatment.water −apparatus
Pollution Control	Water purification	282	+water +treatment −hot −chlorinating −gravimetric −ion −ionization −carbonate −phosphate −ultraviolet −swimming −spas −demineralising −softening −bilge −sterilization −tanks −medical −preparation −substance −agent −filter
Pollution Control	Air purification	240	+toxic.gas
Pollution Control	Water purification	275	+wastewater +treatment −tanks
Waste Management	Process waste	273	+waste +treatment −tanks −repair.animal.waste

Company 2: EnergyHub, Inc. EnergyHub, Inc. applied for this trademark on January 13, 2017, and registered it on March 20, 2018. The Word Mark is "FIRM LOAD DISPATCH" with the registration number 5430216.



STATUS | DOCUMENTS | MAINTENANCE | Download | Print Preview

Generated on:	This page was generated by TSDR on 2026-01-28 22:09:20 EST
Mark:	FIRM LOAD DISPATCH

FIRM LOAD DISPATCH

US Serial Number:	87300454	Application Filing Date:	Jan. 13, 2017
US Registration Number:	5430216	Registration Date:	Mar. 20, 2018
Filed as TEAS RF:	Yes	Currently TEAS RF:	Yes
Register:	Supplemental		
Mark Type:	Service Mark		
TM5 Common Status Descriptor:	LIVE/REGISTRATION/Issued and Active The trademark application has been registered with the Office.		
Amended to Principal Register:	No	Date Amended to Current Register:	Jan. 19, 2018
Status:	A Section 8 declaration has been accepted.		
Status Date:	Dec. 04, 2023		

Mark Information — Expand All

Goods and Services

Note:
The following symbols indicate that the registrant/owner has amended the goods/services:
- Brackets [..] indicate deleted goods/services;
- Double parenthesis ((..)) identify any goods/services not claimed in a Section 15 affidavit of incontestability; and
- Asterisks *..* identify additional (new) wording in the goods/services.

For:	Providing temporary access to non-downloadable software for control of energy management equipment and thermostats; providing temporary use of non-downloadable software for analysis and management of electricity consumption by connected energy management devices		
International Class(es):	042 - Primary Class	U.S Class(es):	100, 101
Class Status:	ACTIVE		
First Use:	Jul. 31, 2016	Use in Commerce:	Jul. 31, 2016

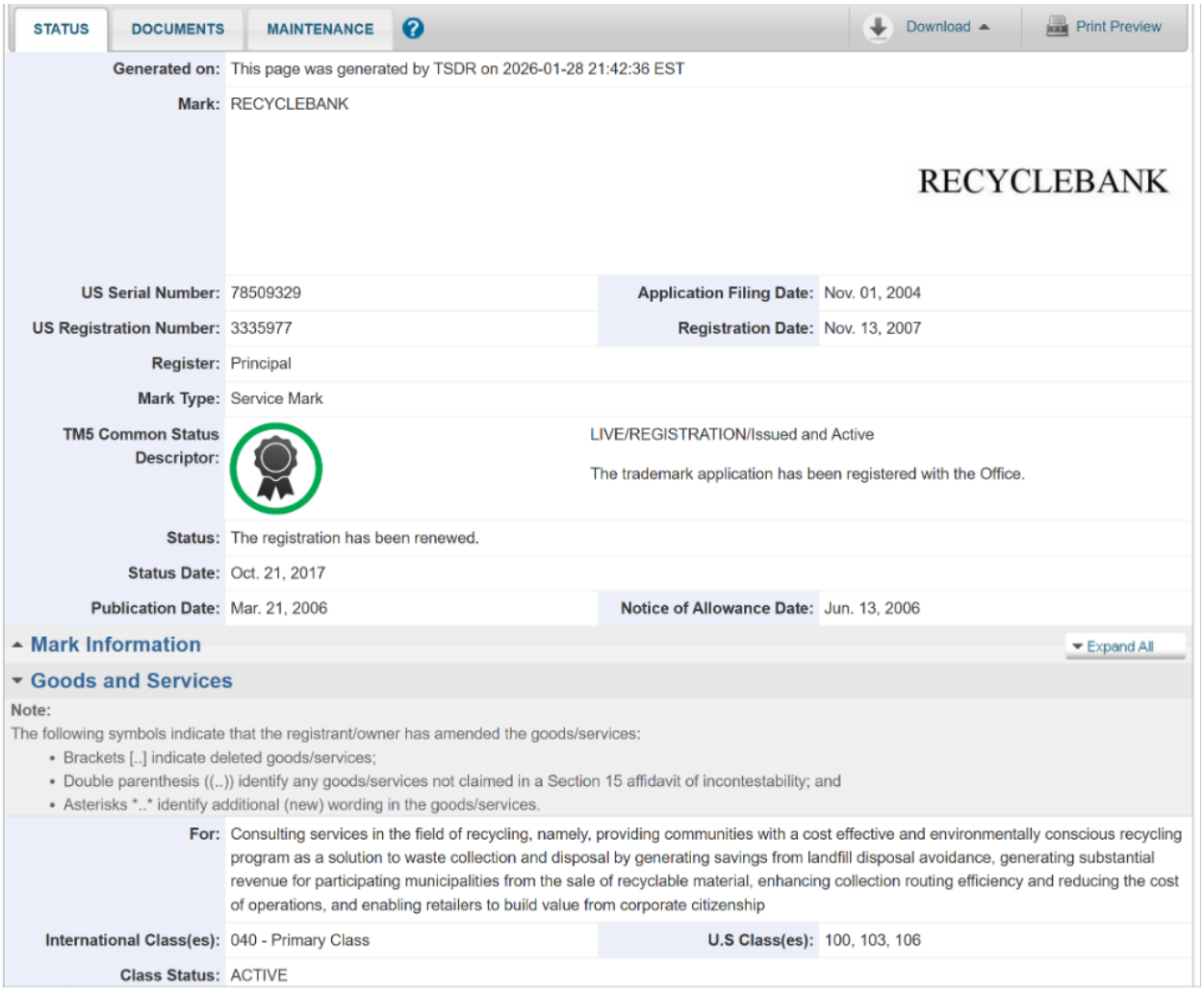
In this example, two terms in the goods and services identification are extracted and shown below. We match each term with any of the 375 green expressions. Both terms are matched to Ref. 88 as follows.

Registration number (Registration date)	Nice Class	Green Expression Ref.	Mark Word	G&S identification at the term level
5430216 (20-Mar-2018)	42	88	FIRM LOAD DISPATCH	Providing temporary access to non-downloadable software for control of energy management equipment and thermostats;
5430216 (20-Mar-2018)	42	88	FIRM LOAD DISPATCH	Providing temporary use of non-downloadable software for analysis and management of electricity consumption by connected energy management devices;

These green terms are matched to Ref. 88, shown below,

Group	Category	Ref.	Green expression
Energy Conservation	Energy management	88	+energy.management

Company 3: RECYCLEBANK LLC RECYCLEBANK LLC applied for this trademark on November 1, 2004, and registered it on November 13, 2007. The Word Mark is "RECYCLEBANK" with the registration number 3335977.



STATUS | DOCUMENTS | MAINTENANCE | Download | Print Preview

Generated on:	This page was generated by TSDR on 2026-01-28 21:42:36 EST
Mark:	RECYCLEBANK

RECYCLEBANK

US Serial Number:	78509329	**Application Filing Date:**	Nov. 01, 2004
US Registration Number:	3335977	**Registration Date:**	Nov. 13, 2007
Register:	Principal		
Mark Type:	Service Mark		
TM5 Common Status Descriptor:	LIVE/REGISTRATION/Issued and Active The trademark application has been registered with the Office.		
Status:	The registration has been renewed.		
Status Date:	Oct. 21, 2017		
Publication Date:	Mar. 21, 2006	**Notice of Allowance Date:**	Jun. 13, 2006

Mark Information Expand All

Goods and Services

Note:
The following symbols indicate that the registrant/owner has amended the goods/services:
- Brackets [..] indicate deleted goods/services;
- Double parenthesis ((..)) identify any goods/services not claimed in a Section 15 affidavit of incontestability; and
- Asterisks *..* identify additional (new) wording in the goods/services.

For:	Consulting services in the field of recycling, namely, providing communities with a cost effective and environmentally conscious recycling program as a solution to waste collection and disposal by generating savings from landfill disposal avoidance, generating substantial revenue for participating municipalities from the sale of recyclable material, enhancing collection routing efficiency and reducing the cost of operations, and enabling retailers to build value from corporate citizenship		
International Class(es):	040 - Primary Class	**U.S Class(es):**	100, 103, 106
Class Status:	ACTIVE		

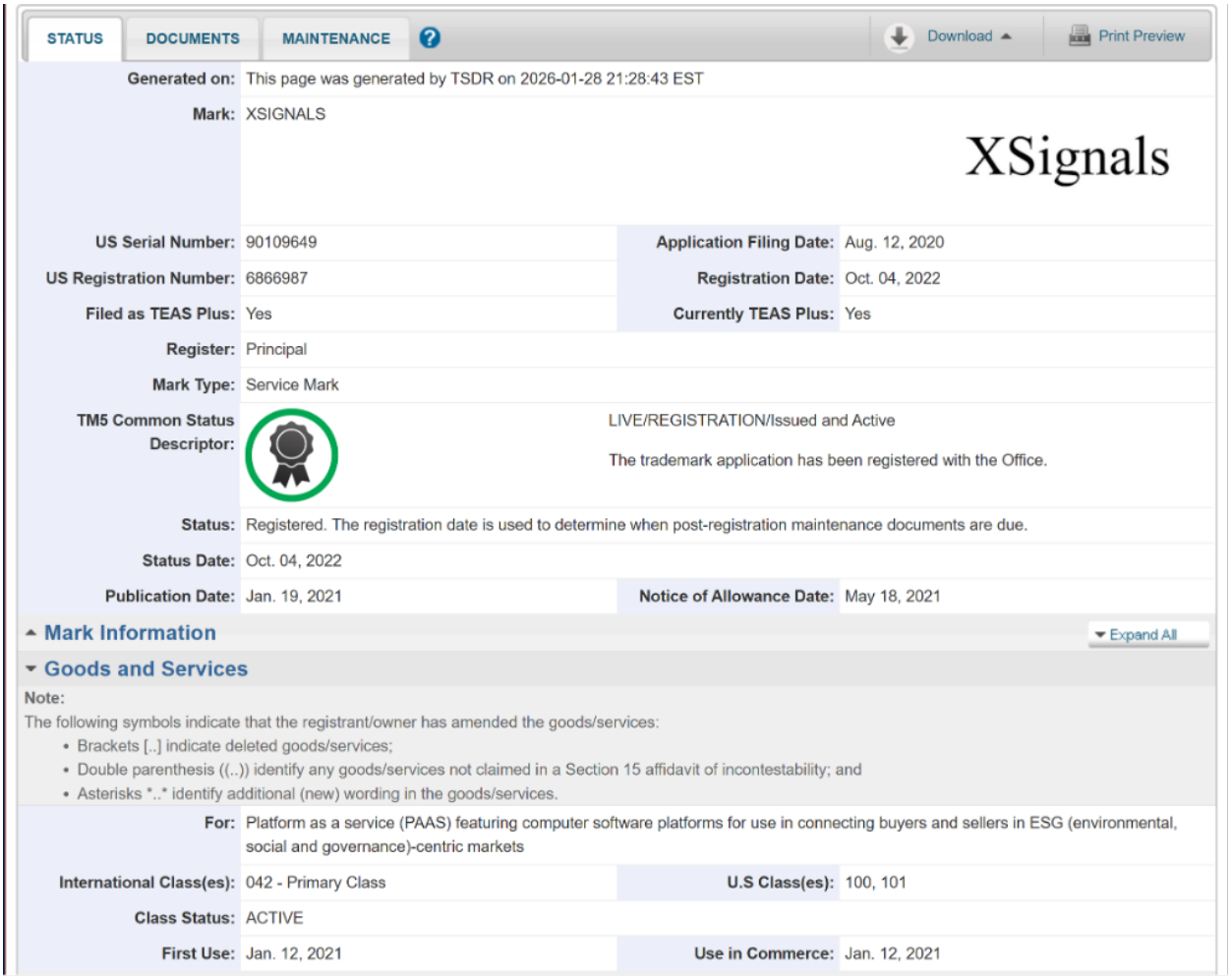
In this example, one term in the goods and services identification are extracted and shown below. We match the term with any of the 375 green expressions. The term is matched to Ref. 97 as follows.

Registration number (Registration date)	Nice Class	Green Expression Ref.	Mark Word	G&S identification at the term level
3335977 (13-Nov-2007)	40	97	RECYCLEBANK	Consulting services in the field of recycling, namely, providing communities with a cost effective and environmentally conscious recycling program as a solution to waste collection and disposal by generating savings from landfill disposal avoidance, generating substantial revenue for participating municipalities from the sale of recyclable material, enhancing collection routing efficiency and reducing the cost of operations, and enabling retailers to build value from corporate citizenship;

At the mark level, this mark is classified as "green" because its G&S identification contains at least one green term, regardless of other non-green terms included. The taxonomy assigns the green term to the group "Climate Change" and the category "Environmental services."

Group	Category	Ref.	Green expression
Climate Change	Environmental services	97	+environmental +services

Company 4: Xpansiv Data Systems, Inc. Xpansiv Data Systems, Inc. applied for this trademark on August 12, 2020, and registered it on October 4, 2022. The Word Mark is "XSignals" with the registration number 6866987.



STATUS | DOCUMENTS | MAINTENANCE | Download | Print Preview

Generated on:	This page was generated by TSDR on 2026-01-28 21:28:43 EST
Mark:	XSIGNALS
	XSignals
US Serial Number:	90109649
Application Filing Date:	Aug. 12, 2020
US Registration Number:	6866987
Registration Date:	Oct. 04, 2022
Filed as TEAS Plus:	Yes
Currently TEAS Plus:	Yes
Register:	Principal
Mark Type:	Service Mark
TM5 Common Status Descriptor:	LIVE/REGISTRATION/Issued and Active The trademark application has been registered with the Office.
Status:	Registered. The registration date is used to determine when post-registration maintenance documents are due.
Status Date:	Oct. 04, 2022
Publication Date:	Jan. 19, 2021
Notice of Allowance Date:	May 18, 2021

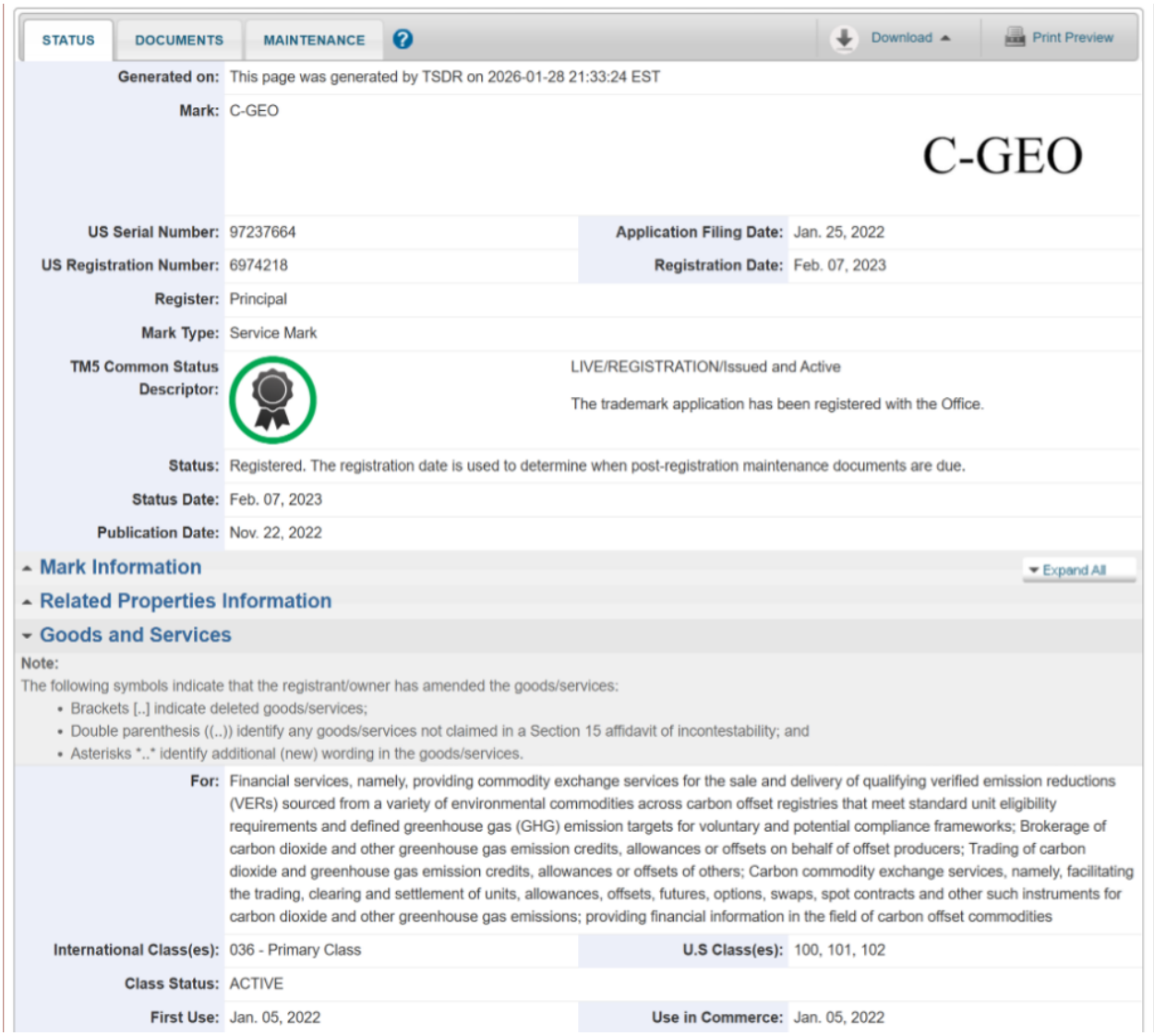
Mark Information — Expand All

Goods and Services

Note:
The following symbols indicate that the registrant/owner has amended the goods/services:
- Brackets [..] indicate deleted goods/services;
- Double parenthesis ((..)) identify any goods/services not claimed in a Section 15 affidavit of incontestability; and
- Asterisks "..." identify additional (new) wording in the goods/services.

For:	Platform as a service (PAAS) featuring computer software platforms for use in connecting buyers and sellers in ESG (environmental, social and governance)-centric markets
International Class(es):	042 - Primary Class
U.S Class(es):	100, 101
Class Status:	ACTIVE
First Use:	Jan. 12, 2021
Use in Commerce:	Jan. 12, 2021

Another example, Xpansiv Data Systems, Inc. applied for this trademark on January 25, 2022, and registered it on February 7, 2023. The Word Mark is "C-GEO" with the registration number 6974218.



STATUS | DOCUMENTS | MAINTENANCE | Download | Print Preview

Generated on:	This page was generated by TSDR on 2026-01-28 21:33:24 EST
Mark:	C-GEO

C-GEO

US Serial Number:	97237664	**Application Filing Date:**	Jan. 25, 2022
US Registration Number:	6974218	**Registration Date:**	Feb. 07, 2023
Register:	Principal		
Mark Type:	Service Mark		
TM5 Common Status Descriptor:	LIVE/REGISTRATION/Issued and Active The trademark application has been registered with the Office.		
Status:	Registered. The registration date is used to determine when post-registration maintenance documents are due.		
Status Date:	Feb. 07, 2023		
Publication Date:	Nov. 22, 2022		

▲ Mark Information — Expand All

▲ Related Properties Information

▼ Goods and Services

Note:
The following symbols indicate that the registrant/owner has amended the goods/services:

- Brackets [..] indicate deleted goods/services;
- Double parenthesis ((..)) identify any goods/services not claimed in a Section 15 affidavit of incontestability; and
- Asterisks *..* identify additional (new) wording in the goods/services.

For:	Financial services, namely, providing commodity exchange services for the sale and delivery of qualifying verified emission reductions (VERs) sourced from a variety of environmental commodities across carbon offset registries that meet standard unit eligibility requirements and defined greenhouse gas (GHG) emission targets for voluntary and potential compliance frameworks; Brokerage of carbon dioxide and other greenhouse gas emission credits, allowances or offsets on behalf of offset producers; Trading of carbon dioxide and greenhouse gas emission credits, allowances or offsets of others; Carbon commodity exchange services, namely, facilitating the trading, clearing and settlement of units, allowances, offsets, futures, options, swaps, spot contracts and other such instruments for carbon dioxide and other greenhouse gas emissions; providing financial information in the field of carbon offset commodities		
International Class(es):	036 - Primary Class	**U.S Class(es):**	100, 101, 102
Class Status:	ACTIVE		
First Use:	Jan. 05, 2022	**Use in Commerce:**	Jan. 05, 2022

After extracting and matching terms with the green taxonomy, we have the following,

Registration number (Registration date)	Nice Class	Green Expression Ref.	Mark Word	G&S identification at the term level
6866987 (4-Oct-2022)	42	97	XSIGNALS	Platform as a service (PAAS) featuring computer software platforms for use in connecting buyers and sellers in ESG (environmental, social and governance)-centric markets;
6974218 (7-Feb-2023)	36	36	C-GEO	Financial services, namely, providing commodity exchange services for the sale and delivery of qualifying verified emission reductions (VERs) sourced from a variety of environmental commodities across carbon offset registries that meet standard unit eligibility requirements and defined greenhouse gas (GHG) emission targets for voluntary and potential compliance frameworks;
6974218 (7-Feb-2023)	36	36	C-GEO	Brokerage of carbon dioxide and other greenhouse gas emission credits, allowances or offsets on behalf of offset producers;
6974218 (7-Feb-2023)	36	36	C-GEO	Trading of carbon dioxide and greenhouse gas emission credits, allowances or offsets of others;
6974218 (7-Feb-2023)	36	36	C-GEO	Carbon commodity exchange services, namely, facilitating the trading, clearing and settlement of units, allowances, offsets, futures, options, swaps, spot contracts and other such instruments for carbon dioxide and other greenhouse gas emissions;
6974218 (7-Feb-2023)	36	36	C-GEO	providing financial information in the field of carbon offset commodities;

Some of the terms are matched to Ref. 36 and 97 and shown below,

Group	Category	Ref.	Green expression
Climate Change	Carbon brokerage	36	+carbon +offsetting
Climate Change	Environmental services	97	+environmental +services

A.1.4 Word clouds by green group

We implement the procedure described in the above section to all marks registered by companies at the intersection of the USPTO trademarks owners and PitchBook companies over the period 1980-2023. The sample comprises 31,299 green terms associated with 14,605 green trademarks.

Figure C.2 plots a word cloud for each group based on these green terms. The font size of each word (or phrase) is proportional to its frequency of occurrence in the green-term corpus. We observe that the top words (phrases) associated with each green group seem to pass face validity. For example, the top phrases for the group Environmental Awareness are “climate change”, “emission reduction”; the top phrases for the group Energy Production are “solar energy”, “renewable energy”;

the top phrases for the group Transportation are "electric vehicle", "vehicle charging"; the top phrases for the group Waste Management are "waste management", "water treatment".

A.1.5 Trends and composition of green products

Table B.1 presents the frequency distribution of green marks across the nine groups defined by the EUIPO Green Trademark Taxonomy. Activity is most heavily concentrated in Energy Production (21.5%), Climate Change (17.9%), and Energy Conservation (16.8%).

Figure C.3 presents the frequency distribution of green trademarks by subperiod. We find that green trademark registrations increase substantially from 1980 to 2020 across all groups, with the largest growth in Energy Production, Climate Change, and Energy Conservation. We also single out the most recent three years (2021–2023). These same groups remain the primary contributors to green products and services in the U.S. market.

A.2 Classifying Green Patents

We draw upon three different classification methodologies from the literature. The "Green Patent" classification described in the paper is the union of these three methodologies, meaning a patent considered as green under *any* methodology is considered a Green Patent.

A.2.1 Green: BKW Classification

We use the mapping of CPC codes to CO_2-reduction technologies developed in Bolton, Kacperczyk, and Wiedemann (2022). This mapping draws upon four distinct sources:

- OECD classification of environmental technologies related to environmental pollution, water scarcity, and climate change mitigation. For details, see Haščič and Migotto (2015).
- Efficiency improving fossil fuel technology classes, from Lanzi, Verdolini, and Haščič (2011).
- Environmentally Sound Technologies, as collected in the IPC Green Inventory (see below).
- Corporate Knights Clean 200, a third-party classification of the top 200 companies by revenue aligned with the Corporate Knights Clean Economy Taxonomy. The BKW classification includes CPC codes found among the patents owned by these firms.

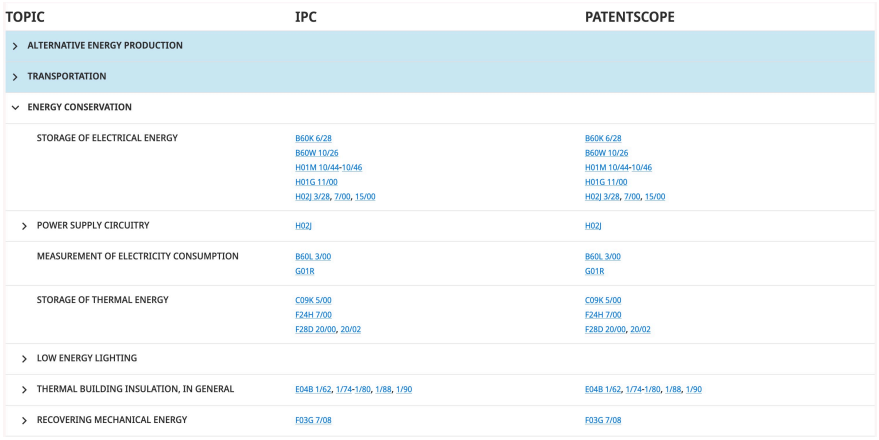
A.2.2 Green: WIPO Classification

The "IPC Green Inventory" is a collection of Environmentally Sound Technologies put together by the IPC Committee of Experts. Figure A.13 contains a screenshot of a sub-sample of the the Inventory.

Starting with the USPTO patent data, we first map USPC to CPC codes using the statistical mapping table provided by the USPTO. We then match those CPC codes using the standard CPC to IPC Concordance. Finally, we identify "Green: WIPO" patents as those for which at least one IPC associated with the patent apears in the IPC Green Inventory.

A.2.3 Green: Robust Classification

We identify "Green: Robust" patents based on the broad CPC category. Specifically, we use CPC codes identified in Table A.2. For additional details, refer to the Github repository for Sun, Tian, and Yang (2024).



TOPIC	IPC	PATENTSCOPE
› ALTERNATIVE ENERGY PRODUCTION		
› TRANSPORTATION		
⌄ ENERGY CONSERVATION		
STORAGE OF ELECTRICAL ENERGY	B60K 6/28 B60W 10/26 H01M 10/44-10/46 H01G 11/00 H02J 3/28, 7/00, 15/00	B60K 6/28 B60W 10/26 H01M 10/44-10/46 H01G 11/00 H02J 3/28, 7/00, 15/00
› POWER SUPPLY CIRCUITRY	H02J	H02J
MEASUREMENT OF ELECTRICITY CONSUMPTION	B60L 3/00 G01R	B60L 3/00 G01R
STORAGE OF THERMAL ENERGY	C09K 5/00 F24H 7/00 F28D 20/00, 20/02	C09K 5/00 F24H 7/00 F28D 20/00, 20/02
› LOW ENERGY LIGHTING		
› THERMAL BUILDING INSULATION, IN GENERAL	E04B 1/62, 1/74-1/80, 1/88, 1/90	E04B 1/62, 1/74-1/80, 1/88, 1/90
› RECOVERING MECHANICAL ENERGY	F03G 7/08	F03G 7/08

Figure A.13: Screenshot of the WIPO Green Inventory Classification

CPC Class	Description
Y02A	Technologies for adaption to Climate Change
Y02B	Climate Change mitigation technologies related to buildings
Y02C	Capture, Storage, Sequestration or Disposal of Greenhouse Gases
Y02D	Climate Change mitigation technologies in information and communication technologies
Y02E	Reduction of Greenhouse Gas emissions, related to energy generation, transmission, or distribution
Y02P	Climate Change mitigation technologies in the production or processing of goods
Y02T	Climate Change mitigation technologies related to transportation
Y02W	Climate Change mitigation technologies related to wastewater treatment or waste management
Y04S	Systems integrating technologies related to power network operation, communication or information technologies for improving the electrical power generation, transmission, distribution, management or usage

Table A.2: CPC Classes used to identify "Green: Robust" patents.

A.3 Green Research Institutes

We collect data on Green Research Institutes by starting with a large list of US Universities and using AI tools to identify green research institutes affiliated with those universities. We begin with a list of 1,726 universities from Unirank. To be included, the universities must be US-based, accredited institutions offering at least three-year undergraduate degrees and/or postgraduate degrees through traditional face-to-face learning in on-site facilities. Two-year degree institutions and distance learning-only institutions are excluded. Both public and private universities are included.

We focus on the top 200 universities, sorted by Unirank's University Ranking algorithm. We view the exact methodology of the ranking, which is derived from a combination of web metrics, as secondary to our analysis; our interest is in collecting a list of green research institutes.

For each of the top 200 universities, we use the OpenAI API to gather information on all green research institutes affiliated with the university. Our system prompt is:

System Prompt

```
You are a careful research assistant. Your task is to identify environmental,
climate, and sustainability-focused academic units at U.S. universities and
extract verifiable founding information. Return your answer in a structure
that matches the provided Pydantic schema:
- ExtractedInstitutes(institutes: List[QueryResponse])
- QueryResponse fields: name, year_founded, confidence, url

Rules:
- Do not guess. If the founding year is not explicitly supported by a reliable
source, set year_founded = null
- Prefer primary sources: the unit's official website ("About", "History"),
official university news/press releases, or governing documents (board/regents
minutes). Avoid Wikipedia and low-quality directories unless they point to
primary sources.
- Include centers, institutes, and schools/colleges/divisions whose mission is
primarily climate/environment/sustainability. Exclude student clubs,
one-off events, and degree programs that are not organizational units.
- If a unit is a rename/reorganization/merger, use the earliest year that is
explicitly supported for the current unit; otherwise set year_founded = null.

Confidence guidance (0.0--1.0):
- 0.9--1.0: founding year explicitly stated on an official unit/university page.
- 0.6--0.8: strong but indirect (e.g., official press release announcing
establishment).
- 0.0--0.5: ambiguous/weak; do not assign a year unless truly supported.

URL:
- Provide the most official canonical webpage for the unit (prefer the unit's
homepage or "About" page). If you cannot find an official page, set url = null.

Important:
- Output ONLY the structured content required by the schema. Do not wrap the
answer in Markdown or add extra commentary.
```

In total, this methodology identifies 878 green research institutes. We further require that the reported confidence score is above 0.9 and the founding year lies between 1980 and 2023, leaving

272 green research institutes affiliated with 150 unique universities.

We label a county as 'treated' once a green research institute opens at a school located in the commuting zone. Thus, we have one observation per commuting zone which corresponds to the earliest green research institute in that commuting zone. This procedure yields a sample of 82 commuting zone-level treatments, reported in Table A.3.

Table A.3: Commuting zones with local green research institutes

Commuting Zone	State	Earliest treatment year	Type of green research institute
Alameda County	California	2007	Institute
Albany County	Wyoming	1993	School/College
Albemarle County	Virginia	1980	Institute
Allegheny County	Pennsylvania	2003	Center
Athens County	Ohio	2004	Institute
Baltimore County	Maryland	2001	Center
Bergen County	New Jersey	1987	Center
Berkshire County	Massachusetts	2024	Center
Bernalillo County	New Mexico	2020	Center
Bexar County	Texas	2010	Institute
Centre County	Pennsylvania	2004	School/College
Champaign County	Illinois	2008	Institute
Chittenden County	Vermont	2017	Institute
Clarke County	Georgia	2007	School/College
Cook County	Illinois	2008	Institute
Cuyahoga County	Ohio	2008	Institute
Dallas County	Texas	2020	Center
Davidson County	Tennessee	1995	Center
Denton County	Texas	1989	Center
Denver County	Colorado	1992	Center
Dyer County	Tennessee	2018	Center
East Baton Rouge Parish	Louisiana	1982	Center
El Paso County	Texas	1990	Center
Erie County	New York	2006	Center
Fairbanks North Star Borough	Alaska	2010	Center
Fairfax County	Virginia	1995	Center
Fayette County	Kentucky	2000	Institute
Forsyth County	North Carolina	2010	Center
Franklin County	Ohio	2019	Institute
Fulton County	Georgia	2004	Institute
Grafton County	New Hampshire	2016	Institute
Hampden County	Massachusetts	2001	Center
Harris County	Texas	1991	Institute
Hennepin County	Minnesota	1991	Institute
Henrico County	Virginia	2022	Institute
Hillsborough County	New Hampshire	1985	Institute
Hillsborough County	Florida	2009	School/College
Houghton County	Michigan	2019	School/College
Jackson County	Missouri	1996	Center
Kalamazoo County	Michigan	2012	Center
King County	Washington	1989	Institute
Kings County	New York	1994	Center
Knox County	Tennessee	2005	Institute

Continued on next page

Commuting Zone	State	Earliest treatment year	Type of green research institute
Lane County	Oregon	2007	Institute
Larimer County	Colorado	1994	Institute
Leon County	Florida	1996	Center
Linn County	Iowa	1987	Center
Los Angeles County	California	1992	Center
Lowndes County	Mississippi	1988	Center
Lucas County	Ohio	2007	School/College
Maricopa County	Arizona	2004	Institute
Marion County	Indiana	1997	Center
Marion County	Florida	2018	Institute
Mecklenburg County	North Carolina	2008	Center
Miami-Dade County	Florida	2016	Center
Middlesex County	Massachusetts	1992	Center
Monroe County	New York	2006	Institute
Monroe County	Indiana	2017	Institute
Multnomah County	Oregon	2009	Center
New Castle County	Delaware	1984	Center
New Haven County	Connecticut	1998	Center
Oklahoma County	Oklahoma	2012	Center
Onondaga County	New York	1993	Center
Palm Beach County	Florida	1994	Center
Philadelphia County	Pennsylvania	1994	Institute
Polk County	Iowa	1987	Center
Providence County	Rhode Island	1998	Institute
Roanoke city	Virginia	2010	Center
Sacramento County	California	1997	Institute
Salt Lake County	Utah	2009	Center
San Diego County	California	2009	Center
Santa Barbara County	California	1991	School/College
Santa Clara County	California	2004	Institute
Shawnee County	Kansas	2003	Center
St. Louis County	Missouri	2023	Center
Tippecanoe County	Indiana	1982	Center
Tuscaloosa County	Alabama	2013	Center
Virginia Beach city	Virginia	2021	Institute
Wake County	North Carolina	1987	Center
Washtenaw County	Michigan	1991	Center
Whitman County	Washington	1991	Center
Yavapai County	Arizona	1992	Institute

B Additional Tables

Table B.1: Green Products by green group and category.

Group	#Green Terms	% Group	Category	#Green Terms	% Category (within group)
Energy Production	6,729	21.50%	Solar Energy	3,052	45.36%
			Other Energy	2,493	37.05%
			Biofuels	827	12.29%
			Wind Energy	357	5.31%
Climate Change	5,599	17.89%	Environmental Services	5,011	89.50%
			Carbon Brokerage	369	6.59%
			Carbon Monitor	219	3.91%
Energy Conservation	5,253	16.78%	Storage Of Electricity	2,262	43.06%
			Energy Saving	1,608	30.61%
			Energy Management	1,156	22.01%
			Low Energy Lighting	227	4.32%
Pollution Control	4,316	13.79%	Water Purification	2,035	47.15%
			Pollution General	1,483	34.36%
			Air Purification	490	11.35%
			Biodegradable	308	7.14%
Environmental Awareness	2,604	8.32%	Ecology	1,409	54.11%
			Sustainability	1,195	45.89%
Reusable	2,535	8.10%	Recycling	2,111	83.27%
			Reusable Bottles	134	5.29%
			Reusable Bags	133	5.25%
			Other Reusable	133	5.25%
			Refilling Cartridge	24	0.95%
Waste Management	2,443	7.81%	Process Waste	2,089	85.51%
			Waste Disposal	354	14.49%
Transportation	1,534	4.90%	General Transport	756	49.28%
			Electric Engines	445	29.01%
			Electric Bike	96	6.26%
			Other Vehicles	90	5.87%
			Electric Moto	59	3.85%
			Hydrogen Vehicle	31	2.02%
			Hybrid Vehicle	30	1.96%
			Electric Car	27	1.76%
Agriculture	286	0.91%	Fertiliser Alternatives	244	85.31%
			Pesticide Alternatives	35	12.24%
			Other Agriculture	7	2.45%

This table reports the frequency distribution of green terms by group. The sample comprises 31,299 green terms associated with 14,605 green trademarks registered by corporations with the USPTO during 1980–2023. The EUIPO Green Trademark Taxonomy classifies these terms into nine groups: (1) Agriculture, (2) Climate Change, (3) Energy Conservation, (4) Energy Production, (5) Environmental Awareness, (6) Pollution Control, (7) Reuse and Reusability, (8) Transportation, and (9) Waste Management.

Table B.2: American Recovery and Reinvestment Act Spending, by Awarding Agency

Awarding Agency Name	Aggregate Award for Non-Green Spending	Aggregate Award for Green Spending
Administration for Children and Families	5,165,000,000	
Administration on Aging	124,800,000	
Agency for Healthcare Research and Quality	546,100,000	
Agricultural Marketing Service	67,506,611	
Agricultural Research Service	177,100,000	
Air Force Center for Studies and Analyses	1,647,994	
Air Force Management Engineering Agency	4,892,486	
Air Force Materiel Command	11,339,074	
Air Force Operational Test and Evaluation Center	3,556,792	
Air Mobility Command	13,233,880	
Air National Guard	47,926,558	
Air National Guard Units (Mobilization) (Title 5)	733,682	
Air Training Command	11,487,715	
Ames Research Center	805,089	
Argonne Site Office		66,200
Army National Guard Units (Title 32)	6,553,320	
Assistant Administrator for Administration and Resources Management		699,904
Assistant Administrator for Solid Waste and Emergency Response		8,500,000
Assistant Secretary for Administration	354,550	
Assistant Secretary for Community Planning and Development	511,900,000	
Assistant Secretary for Energy Efficiency and Renewable Energy		20,768,591
Assistant Secretary for Environmental Management		2,135,821
Assistant Secretary for Fair Housing and Equal Opportunity	811,117	
Assistant Secretary for Fossil Energy		211,400,000
Assistant Secretary for Information and Technology	5,749,309	
Assistant Secretary for Public and Indian Housing	872,400,000	
Bureau of Alcohol, Tobacco, Firearms and Explosives	6,231,877	
Bureau of Land Management	224,800,000	
Bureau of Reclamation	879,900,000	

Continued on next page

Awarding Agency Name	Aggregate Award for Non-Green Spending	Aggregate Award for Green Spending
Bureau of the Census	135,200,000	
Bureau of the Public Debt	1,605,384	
Carlsbad Field Office		40,000
Centers for Disease Control and Prevention	1,100,000,000	
Centers for Medicare & Medicaid Services	392,600,000	
Chicago Office		150,000
Community Development Financial Institutions	98,169,266	
Congress, generally, no additional specification available	4,284,781	
Consolidated Business Center		747,019
Cooperative State Research, Education, and Extension Service	34,098,862	
Corporation for National and Community Service	119,900,000	
Defense Logistics Agency	43,898,340	
Department of Agriculture	18,099,831	
Department of Commerce	126,824	
Department of Defense (except military departments)	3,547,649	
Department of Education	2,953,000,000	
Department of Energy		40,040,000,000
Department of Health and Human Services	1,856,000,000	
Department of Homeland Security	9,751,545	
Department of Housing and Urban Development	10,490,000,000	
Department of Justice	4,107,000,000	
Department of Labor	6,236,666	
Department of State	223,100,000	
Department of Transportation	235,600,000	
Department of Veterans Affairs	1,422,000,000	
Department of the Air Force	1,145,000,000	
Department of the Army	1,556,000,000	
Department of the Interior	20,560,818	
Department of the Navy	1,743,000,000	
Deputy Secretary of Housing and Urban Development	273,875	
Director, OMB	13,574	
Dwight D. Eisenhower Memorial Commission	8,824	
Economic Development Administration	145,700,000	
Employment Standards Administration	9,383	
Employment and Training Administration	4,496,000,000	

Continued on next page

Awarding Agency Name	Aggregate Award for Non-Green Spending	Aggregate Award for Green Spending
Environmental Protection Agency		7,185,000,000
Export-Import Bank of the U.S.	49,715	
Farm Service Agency	104,000,000	
Federal Aviation Administration	1,364,000,000	
Federal Communications Commission	96,369,384	
Federal Emergency Management Agency	645,200,000	
Federal Energy Regulatory Commission		24,331,431
Federal Highway Administration	27,920,000,000	
Federal Mediation and Conciliation Service	5,457,856	
Federal Railroad Administration	9,458,000,000	
Federal Student Aid	20,747,405	
Federal Transit Administration	9,628,000,000	
Fish and Wildlife and Parks (Assistant Secretary)	420,829	
Food and Drug Administration	12,239,356	
Food and Nutrition Service	297,100,000	
Forest Service	1,159,000,000	
GAO, except Comptroller General	391,139	
General Services Administration	577,600,000	
Geological Survey	148,300,000	
Golden Field Office		417,200,000
Government Accountability Office	2,601,836	
HUD Board of Contract Appeals	82,364	
Headquarters, Air Force Reserve	1,611,343	
Headquarters, Army Materiel Command	47,456	
Health Resources and Services Administration	2,652,000,000	
Idaho Operations Office		40,200,000
Immediate Office of the Administrator of EPA		960,000
Immediate Office of the Secretary of Health and Human Services	176,093	
Impact Aid Programs	53,029	
Indian Affairs (Assistant Secretary)	935,400,000	
Indian Health Service	584,400,000	
Institute of Education Sciences	12,972,730	
International Institute for Cotton	289,400	
Joint House and Senate Entities	264,696	
Kansas City Site Office		40,308
Maritime Administration	258,000,000	
National Aeronautics and Space Administration	1,097,000,000	
National Business Center	12,858	

Continued on next page

Awarding Agency Name	Aggregate Award for Non-Green Spending	Aggregate Award for Green Spending
National Endowment for the Arts	49,847,122	
National Energy Technology Laboratory		1,495,000,000
National Guard Bureau	5,574,740	
National Highway Traffic Safety Administration	60,235	
National Institute of Standards and Technology	1,374,000,000	
National Institutes of Health	10,610,000,000	
National Nuclear Security Administration Service Center		188,090
National Oceanic and Atmospheric Administration	4,115,000,000	
National Park Service.	746,000,000	
National Railroad Passenger Corporation (AMTRAK)	9,199,339	
National Science Foundation	3,016,000,000	
National Telecommunication and Information Administration	33,972,800	
Natural Resources Conservation Service	242,700,000	
Naval Air Systems Command	455,290	
Naval Facilities Engineering Command	188,100,000	
Naval Sea Systems Command	1,249,994	
Naval Space and Warfare Systems Command	35,649,549	
OSDBU/Minority Business Resources Center	19,000	
Oak Ridge Office		9,636,549
Occupational Safety and Health Administration	1,298,854	
Office Healthy Homes and Lead Hazard Control	99,117,098	
Office for Civil Rights	388,132	
Office of Assistant Secretary for Administration and Management	30,375,910	
Office of Assistant Secretary for Planning and Evaluation	5,786,540	
Office of Chief Financial Officer	148,128	
Office of Civilian Radioactive Waste Management		780,700
Office of Electricity Delivery and Energy Reliability		43,152,336
Office of Elementary and Secondary Education	65,760,000,000	

Continued on next page

Awarding Agency Name	Aggregate Award for Non-Green Spending	Aggregate Award for Green Spending
Office of Field Policy and Management Region VI, Fort Worth Regional Office	102,784	
Office of Field Policy and Management Region VIII, Denver Regional Office	232,439	
Office of Governmentwide Policy	51,000	
Office of Hearings and Appeals		525,692
Office of Higher Education Programs	656,029	
Office of Homeland Security	1,735,398	
Office of Indian Education	179,236	
Office of Innovation and Improvement	682,700,000	
Office of Inspector General	5,018,999	
Office of Intergovernmental Affairs and Regional Directors	31,534,289	
Office of Job Corps	190,334	
Office of Management and Budget	145,500	
Office of Naval Research	2,595,864	
Office of Nuclear Energy, Science and Technology		35,780
Office of Nuclear Security/National Nuclear Security Administration		43,481,000
Office of Policy, Management and Budget/Chief Financial Officer	117,900,000	
Office of Postsecondary Education	267,740	
Office of Science		956,200,000
Office of Special Education and Rehabilitative Services	13,670,000,000	
Office of Surface Mining, Reclamation and Enforcement	154,422	
Office of the Assistant Secretary for Administration and Management	257,800,000	
Office of the Chief People Officer	6,488,377	
Office of the Chief Procurement Officer	1,110,118	
Office of the Federal Acquisition Service	184,900,000	
Office of the General Counsel	50,000	
Office of the Inspector General	787,040	2,010,868
Office of the Secretary	257,268	286,413
Office of the Secretary of Health and Human Services	94,795,963	
Office of the Secretary of Housing and Urban Development	4,593,972	
Office of the Secretary of State	488	
Office of the Secretary of Transportation	2,157,158	

Continued on next page

Awarding Agency Name	Aggregate Award for Non-Green Spending	Aggregate Award for Green Spending
Office of the Secretary, Department of Homeland Security	16,810,656	
Office of the Solicitor	8,300	
Office the of the Energy Information Administration		1,058,381
Offices of the Regional Administrators	6,904,409	
Other Small Agencies	1,721,662	
Overseas Private Investment Corporation	10,786	
President of the United States	120,349	
Program Support Center	342,000,000	
Public Buildings Service	6,256,000,000	
Railroad Retirement Board	58,575	
Recovery Accountability and Transparency Board	569,160	
Rehabilitation Services Administration	201,169	
Rural Business Cooperative Service	19,505,179	
Rural Housing Service	1,525,000,000	
Rural Utilities Service	6,710,000,000	
Savannah River Operations Office		211,907
School Support and Technology Programs	263	
Small Business Administration	109,900,000	
Smithsonian Institution	24,806,164	
Social Security Administration	52,446,506	
Space Command	2,232,971	
State Justice Institute	9,569	
Substance Abuse and Mental Health Services Administration	7,423,951	
Surface Transportation Board - formerly ICC - code 3000	1,403,841	
The Legislative Branch	7,069,967	
Transportation Security Administration	1,054,000,000	
U.S. Agency for International Development	19,979,317	
U.S. Air Force Academy	18,300,000	
U.S. Army Communications Electronics Command	135,000	
U.S. Army Corps of Engineers - civil program financing only	4,889,000,000	
U.S. Army Corps of Engineers, except civil program financing	2,230,000,000	
U.S. Army Medical Command	8,977	
U.S. Army Research Laboratory Command	12,150,000	
U.S. Army Research, Development and Engineering Command	1,454,577	

Continued on next page

Awarding Agency Name	Aggregate Award for Non-Green Spending	Aggregate Award for Green Spending
U.S. Army Soldiers System Command (SSC)	59,164	
U.S. Army Tank-Auto. & Armament Command (TACOM)	8,748,000	
U.S. Coast Guard	310,300,000	
U.S. Customs and Border Protection	433,100,000	
U.S. Fish and Wildlife Service	318,400,000	
U.S. Immigration and Customs Enforcement	20,075,347	
U.S. Postal Inspection Service	62,871	
U.S. Tax Court	180,467	
U.S. and Mexico International Boundary and Water Commission	244,800,000	
USDA, Office of Operations	17,460,586	
USDA, Office of Small and Disadvantaged Business	207,167	
USDA, Office of the Chief Economist	24,335	
USDA, Office of the Chief Financial Officer	7,934,956	
USDA, Office of the Chief Information Officer	10,547,584	
Under Secretary for Health / Veterans Health Administration	26,690,701	
Under Secretary for Rural Development	30,000	
Western Area Power Marketing Administration		3,325,247
White House Commission on the National Moment of Remembrance	975,000	

Notes: Source: NBER American Recovery and Reinvestment Act dataset.

Table B.3: Summary statistics of Green Research Institute startup sample

	N	Mean	SD	p5	p25	Median	p75	p95
Green Research Institute$_{c,t-1}$	118,633	0.564	0.496	0	0	1	1	1
Green Trademarks (ratio) $(t \to t+1)$	17,267	0.032	0.168	0	0	0	0	0
Green Trademarks (ratio) $(t \to t+2)$	26,597	0.035	0.176	0	0	0	0	0
Green Trademarks (ratio) $(t \to t+3)$	32,636	0.037	0.181	0	0	0	0	0
Green Trademarks (ratio) $(t \to t+4)$	37,201	0.039	0.184	0	0	0	0	0.071
Green Trademarks (ratio) $(t \to t+5)$	40,704	0.040	0.186	0	0	0	0	0.143
Green Trademarks (ratio) $(t \to t+6)$	43,374	0.041	0.186	0	0	0	0	0.200
Green Patents (ratio) $(t \to t+1)$	4,392	0.221	0.384	0	0	0	0.323	1
Green Patents (ratio) $(t \to t+2)$	7,548	0.247	0.403	0	0	0	0.500	1
Green Patents (ratio) $(t \to t+3)$	10,510	0.263	0.408	0	0	0	0.500	1
Green Patents (ratio) $(t \to t+4)$	12,660	0.266	0.405	0	0	0	0.500	1
Green Patents (ratio) $(t \to t+5)$	14,215	0.271	0.404	0	0	0	0.500	1
Green Patents (ratio) $(t \to t+6)$	15,300	0.273	0.401	0	0	0	0.500	1

Notes: This table presents summary statistics of the regression sample corresponding to Table 5.

Table B.4: Summary statistics of Green Research Institute commuting zone sample

	N	Mean	SD	p5	p25	Median	p75	p95
Green Research Institute$_{c,t-1}$	4,779	0.212	0.409	0	0	0	0	1
Green-innovation Startups (counts) $t, t+1$	4,779	0.416	1.413	0	0	0	0	2
Green-innovation Startups (counts) $t, t+2$	4,779	0.730	2.499	0	0	0	1	3
Green-innovation Startups (counts) $t, t+3$	4,779	1.051	3.672	0	0	0	1	5
Green-innovation Startups (counts) $t, t+4$	4,779	1.299	4.543	0	0	0	1	5
Green-innovation Startups (counts) $t, t+5$	4,779	1.420	4.954	0	0	0	1	6
Green-innovation Startups (counts) $t, t+6$	4,779	1.506	5.230	0	0	0	1	6

Notes: This table presents summary statistics of the regression sample corresponding to Panel (a) of Table 6.

Table B.5: Summary statistics of Green Research Institute commuting zone sample

	N	Mean	SD	p5	p25	Median	p75	p95
Green Research Institute$_{c,t-1}$	9,309	0.130	0.337	0	0	0	0	1
N^{50}	9,309	6.971	22.790	0	1	1	4	29
N^{75}	9,309	4.058	12.578	0	0	1	3	17
N^{90}	9,309	2.329	7.076	0	0	1	2	10
N^{50}_{ind}	9,309	7.002	24.369	0	1	1	4	30
N^{75}_{ind}	9,309	4.077	13.662	0	0	1	3	17
N^{90}_{ind}	9,309	2.377	7.663	0	0	1	2	10

Notes: This table presents summary statistics of the regression sample corresponding to Panel (b) of Table 6.

Table B.6: Summary statistics of Green ARRA startup sample

	N	Mean	SD	p5	p25	Median	p75	p95
$\log \frac{\text{Green ARRA}}{\text{Population}}$	84,340	4.026	2.226	0	2.638	4.942	5.764	6.676
Post 2009	84,340	0.253	0.435	0	0	0	1	1
Post 2013	84,340	0.421	0.494	0	0	0	1	1
Green Trademarks (ratio) $(t \to t+1)$	13,503	0.033	0.172	0	0	0	0	0
Green Trademarks (ratio) $(t \to t+2)$	20,985	0.037	0.181	0	0	0	0	0
Green Trademarks (ratio) $(t \to t+3)$	25,787	0.039	0.185	0	0	0	0	0
Green Trademarks (ratio) $(t \to t+4)$	29,324	0.042	0.190	0	0	0	0	0.167
Green Trademarks (ratio) $(t \to t+5)$	32,084	0.043	0.191	0	0	0	0	0.250
Green Trademarks (ratio) $(t \to t+6)$	34,151	0.044	0.193	0	0	0	0	0.286
Green Patents (ratio) $(t \to t+1)$	3,475	0.238	0.396	0	0	0	0.368	1
Green Patents (ratio) $(t \to t+2)$	6,078	0.262	0.412	0	0	0	0.500	1
Green Patents (ratio) $(t \to t+3)$	8,486	0.275	0.414	0	0	0	0.571	1
Green Patents (ratio) $(t \to t+4)$	10,207	0.278	0.412	0	0	0	0.571	1
Green Patents (ratio) $(t \to t+5)$	11,474	0.282	0.409	0	0	0	0.600	1
Green Patents (ratio) $(t \to t+6)$	12,320	0.285	0.407	0	0	0	0.600	1

Notes: This table presents summary statistics of the regression sample corresponding to Table 8.

Table B.7: Summary statistics of Green ARRA startup sample

	N	Mean	SD	p5	p25	Median	p75	p95
log Green ARRA	2,924	15.916	5.217	0	14.702	17.085	19.289	20.791
Post 2009	2,924	0.248	0.432	0	0	0	0	1
Post 2013	2,924	0.320	0.466	0	0	0	1	1
Green-innovation Startups (counts) $(t \to t+1)$	2,924	0.560	1.669	0	0	0	0	3
Green-innovation Startups (counts) $(t \to t+2)$	2,924	0.991	2.974	0	0	0	1	5
Green-innovation Startups (counts) $(t \to t+3)$	2,924	1.415	4.359	0	0	0	1	7
Green-innovation Startups (counts) $(t \to t+4)$	2,924	1.747	5.412	0	0	0	1	8
Green-innovation Startups (counts) $(t \to t+5)$	2,924	1.913	5.922	0	0	0	1	9
Green-innovation Startups (counts) $(t \to t+6)$	2,924	2.027	6.267	0	0	0	1	9

Notes: This table presents summary statistics of the regression sample corresponding to Panel (a) of Table 9.

Table B.8: Summary statistics of Green ARRA commuting zone sample

	N	Mean	SD	p5	p25	Median	p75	p95
log Green ARRA	4,792	14.002	6.390	0	13.411	15.441	18.588	20.435
Post 2009	4,792	0.232	0.422	0	0	0	0	1
Post 2013	4,792	0.291	0.454	0	0	0	1	1
N^{50}	4,792	9.309	28.620	0	1	2	6	43
N^{75}	4,792	5.322	15.584	0	0	1	3	25
N^{90}	4,792	3.006	8.606	0	0	1	2	15
N^{50}_{ind}	4,792	9.518	30.787	0	1	2	5	42
N^{75}_{ind}	4,792	5.456	17.007	0	0	1	3	25
N^{90}_{ind}	4,792	3.103	9.385	0	0	1	2	15

Notes: This table presents summary statistics of the regression sample corresponding to Panel (b) of Table 9.

Table B.9: Summary statistics of Natural Disaster startup sample

	N	Mean	SD	p5	p25	Median	p75	p95
$\text{Disaster}_{t-1}^{\text{Adj}}$	419,554	0.468	0.499	0	0	0	1	1
$\log \text{Population}_{t-1}$	360,645	13.396	1.542	10.498	12.578	13.482	14.262	16.106
$\log \text{Income}_{t-1}$	360,645	17.176	1.827	13.757	16.114	17.344	18.398	20.133
Green Trademarks (ratio) $(t \to t+1)$	29,258	0.034	0.173	0	0	0	0	0
Green Trademarks (ratio) $(t \to t+2)$	45,843	0.037	0.178	0	0	0	0	0
Green Trademarks (ratio) $(t \to t+3)$	57,572	0.038	0.179	0	0	0	0	0.100
Green Trademarks (ratio) $(t \to t+4)$	66,653	0.038	0.178	0	0	0	0	0.150
Green Trademarks (ratio) $(t \to t+5)$	73,846	0.039	0.177	0	0	0	0	0.200
Green Trademarks (ratio) $(t \to t+6)$	79,976	0.039	0.177	0	0	0	0	0.200
Green Patents (ratio) $(t \to t+1)$	19,519	0.225	0.381	0	0	0	0.333	1
Green Patents (ratio) $(t \to t+2)$	27,295	0.232	0.381	0	0	0	0.333	1
Green Patents (ratio) $(t \to t+3)$	32,172	0.236	0.381	0	0	0	0.375	1
Green Patents (ratio) $(t \to t+4)$	35,501	0.239	0.380	0	0	0	0.400	1
Green Patents (ratio) $(t \to t+5)$	37,918	0.239	0.378	0	0	0	0.400	1
Green Patents (ratio) $(t \to t+6)$	39,909	0.239	0.376	0	0	0	0.400	1

Notes: This table presents summary statistics of the regression sample corresponding to Table 10.

Table B.10: Summary statistics of Natural Disaster county sample

	N	Mean	SD	p5	p25	Median	p75	p95
$\text{Disaster}_{t-1}^{\text{Adj}}$	53,442	0.147	0.354	0	0	0	0	1
$\log \text{Population}_{t-1}$	37,035	10.056	1.175	8.404	9.314	9.923	10.604	12.319
$\log \text{Income}_{t-1}$	37,035	13.123	1.503	10.981	12.086	12.972	13.954	15.949
Green-innovation Startups (counts) $(t \to t+1)$	53,442	0.007	0.156	0	0	0	0	0
Green-innovation Startups (counts) $(t \to t+2)$	53,442	0.011	0.249	0	0	0	0	0
Green-innovation Startups (counts) $(t \to t+3)$	53,442	0.015	0.347	0	0	0	0	0
Green-innovation Startups (counts) $(t \to t+4)$	53,442	0.018	0.408	0	0	0	0	0
Green-innovation Startups (counts) $(t \to t+5)$	53,442	0.021	0.453	0	0	0	0	0
Green-innovation Startups (counts) $(t \to t+6)$	53,442	0.022	0.485	0	0	0	0	0
N^{50}	53,442	0.180	2.809	0	0	0	0	0
N^{75}	53,442	0.089	1.345	0	0	0	0	0
N^{90}	53,442	0.035	0.564	0	0	0	0	0
N_{ind}^{50}	53,442	0.038	0.675	0	0	0	0	0
N_{ind}^{75}	53,442	0.092	1.559	0	0	0	0	0
N_{ind}^{90}	53,442	0.184	3.097	0	0	0	0	0

Notes: This table presents summary statistics of the regression sample corresponding to Table 11.

C Additional Figures

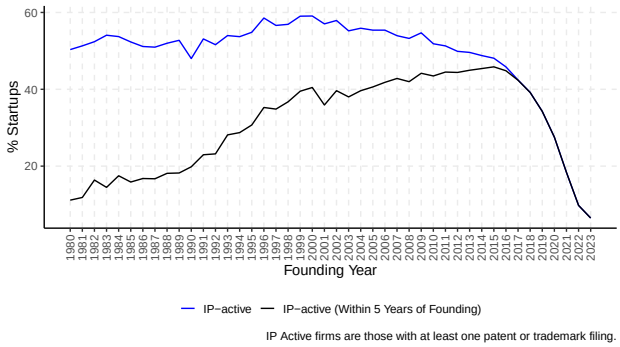


Figure C.1: The fraction of startups founded each year which engage in intellectual property activity.

Notes: The blue line is the fraction of startups founded each year that go on to ever file a patent or trademark. The black line is the fraction of startups founded each year that go on to file a patent or trademark within three years of founding.



Figure C.2: Word clouds by green group. The sample comprises 31,299 green terms associated with 14,605 green trademarks registered by corporations with the USPTO during 1980–2023. The EUIPO Green Trademark Taxonomy classifies these terms into nine groups: (1) Agriculture, (2) Climate Change, (3) Energy Conservation, (4) Energy Production, (5) Environmental Awareness, (6) Pollution Control, (7) Reuse and Reusability, (8) Transportation, and (9) Waste Management. The figure presents word clouds for each group.

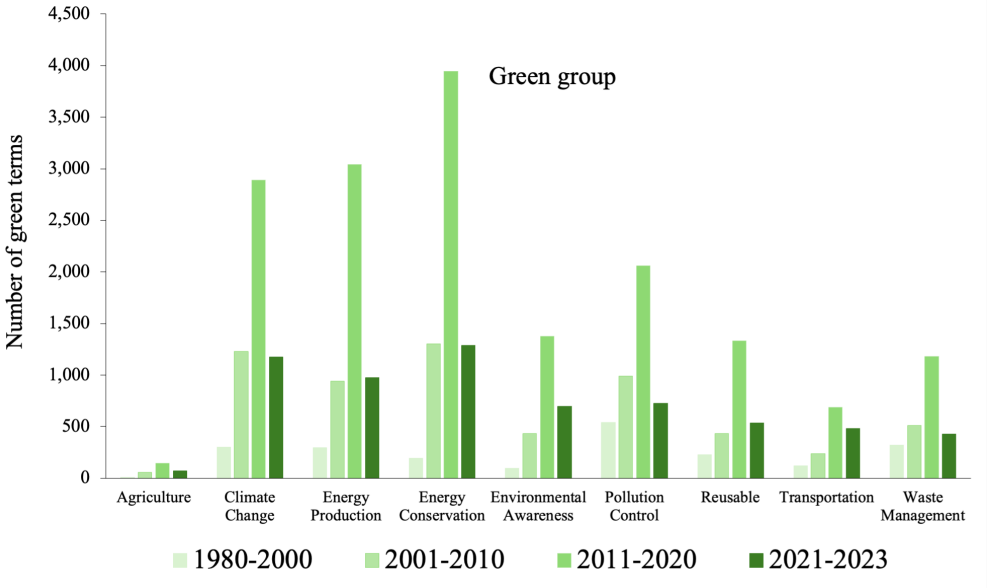


Figure C.3: The sample comprises 31,299 green terms associated with 14,605 green trademarks registered by corporations with the USPTO during 1980–2023. The EUIPO Green Trademark Taxonomy classifies these terms into nine groups: (1) Agriculture, (2) Climate Change, (3) Energy Conservation, (4) Energy Production, (5) Environmental Awareness, (6) Pollution Control, (7) Reuse and Reusability, (8) Transportation, and (9) Waste Management.