Innovative Nonprofits: Evidence from Patents

Elaine Kiiru*

September 22, 2024

Abstract

In endogenous growth theory, patenting is motivated by profits. However, this paper finds that nonprofit organizations own about two percent of granted patents in the United States. Furthermore, nonprofit patents document important contributions to innovation. Text analysis shows that nonprofit patents make disproportionate contributions to technological categories that have positive externalities, such as medical, green, and basic research. High citations and high novelty relative to impact, as measured by the similarity between past and future patents, suggest that nonprofit patents generate notable spillovers. These contributions not only enhance technological and societal impact but also align closely with the nonprofits' prosocial missions.

Keywords: nonprofit, patent

JEL Codes: O33, O34

*I am grateful for the guidance of my advisors: David Lagakos, Stephen Terry, and Masao Fukui. This project was supported by the GivingTuesday Data Commons Fellowship.

1 Introduction

The endogenous growth theory literature, stemming from Romer (1990), assumes that patents align incentives to innovate with firms' profit motive. Firms operating under a profit motive only consider the private value of their innovations, not valuing the societal value of innovations with positive externalities. Yet, there are many types of innovation, from technologies that facilitate further development to environmental and health technologies, that have been shown to have positive externalities. These technologies, and similarly technologies that are public goods, are less likely to be provided by for-profit firms. This paper presents new evidence from patents that nonprofit organizations undertake these important innovations.

I find that over a hundred thousand patents filed since 1990 are assigned to nonprofit organizations. These patents tend to receive higher citations and have high importance scores, based on comparing backwards and forwards textual similarity with other patents, suggesting that nonprofit patents have positive externalities through high spillovers. Nonprofit patents also tend to be concentrated in technological categories that contain public goods or have positive externalities, based on textual analysis of patent abstracts and Cooperative Patent Classification (CPC) class titles using a Large Language Model (LLM). Nonprofit involvement in the production of technologies with positive externalities suggests that patenting nonprofit organizations pursue their prosocial agendas directly through innovation, rather than primarily as a means of funding other projects. However, endogenous growth theory, and policies based on it, overlook the altruistic motivations of the nonprofit sector by focusing on profit as the motive for innovation.

This paper is the first to identify and describe nonprofit patenting in a way that distinguishes between an organization's function and legal structure. I achieve this by matching patent assignee names to nonprofit organization names using Levenshtein ratios of string similarity. I find that about half of nonprofit patents are assigned to universities, while twenty percent belong to health organizations and public and societal benefit organizations respectively. Although nonprofit patents only account for about two percent of all assigned patents between 1989 and 2017, high quality data on the characteristics of the innovations that they represent permit a detailed comparison between them and patents owned by other types of organizations.

Nonprofit patents are concentrated in categories of technology that are likely to be under provided by the market because they have positive externalities or are public goods. About 60% of nonprofit patenting happens within CPC classes pertaining to "Biochemistry; ... ; Mutation Or Genetic Engineering", "Organic Chemistry", "Medical Or Veterinary Science; Hygiene", or "Measuring; Testing". These categories reflect types of technologies that literature has shown to be difficult for the market to provide at a socially optimal level due to externalities. Examples of these include health technologies, green technologies, and basic research (Acemoglu et al., 2012; Akcigit, Hanley, and Serrano-Velarde, 2021; J.B. Rebitzer and R.S. Rebitzer, 2023). Using a LLM to analyze patent abstracts, I find that a higher percentage of patents from nonprofit organizations fall into these categories. I also use a LLM to test the names of CPC classes, themselves, showing that CPC classes with high nonprofit patenting tend to contain public goods. I manually validate the results of my LLM.

Nonprofit patents also provide positive externalities through spillovers. Akcigit, Hanley, and Serrano-Velarde (2021) show that basic research has bigger spillovers than applied research. While pure basic research is not patentable due to the utility requirement, use-inspired basic research is more common and pure applied research is less common among nonprofit patents. Higher citations provide evidence that nonprofit patents are highly effective at disseminating information that is used by other inventors. These results hold even when controlling for CPC class, year, and the purpose of the organization. Strong representation among important patents from the last thirty years provides evidence that nonprofit patents are more innovative than patents owned by other entities. Patents that represent larger advances in technology (Kelly et al., 2021) and have greater spillovers (Akcigit, Hanley, and Serrano-Velarde, 2021) provide greater productivity improvements, advancing economic growth.

My findings complement other papers that distinguish between different assignee organization characteristics. While several papers match patents to their for-profit assignees (Hall, A.B. Jaffe, and Trajtenberg, 2001; Kogan et al., 2017; Autor et al., 2020), I am the first to provide a crosswalk between nonprofit organizations and patents. Akcigit, Hanley, and Serrano-Velarde (2021) find that patents assigned to government entities receive more citations, which implies that they produce technologies with a larger step size. In Trajtenberg, Henderson, and A. Jaffe (1997), university patents tend to be more important and general than corporate patents based on citations. I show that belonging to a nonprofit organization matters independently from belonging to a university.

This paper contributes to a literature that links patent assignee characteristics with patent quality. For public for-profit firms, stock prices provide one way to evaluate the quality of patents (Kogan et al., 2017; Kelly et al., 2021; Kalyani, 2022). Kelly et al. (2021), Krieger, Schnitzer, and Watzinger (2022), and Kalyani (2022) analyze the text of patent abstracts to determine their quality, novelty, creativity, or importance. Other papers base their measurements of patent quality on citations (Hall, A.B. Jaffe, and Trajtenberg, 2001; Akcigit and Kerr, 2018; Lerner and Seru, 2022). In addition to these methods, I use a LLM for zero-shot classification, to evaluate qualities of nonprofit patents in comparison with patents owned by other types of assignees.

Nonprofits may help correct distortions in the direction of innovation by investing in the most socially beneficial technologies with the highest spillovers. The technological path dependence literature suggests that in the presence of spillovers and externalities, and markups, directed subsidies can push innovation towards a more socially optimal outcome (Acemoglu, 2023; Donald, 2023). Hence, in the presence of information or oversight constraints about research projects undertaken, the government or private donors could improve the efficiency of research funding by directing research funds to nonprofit organizations.

The next section describes the creation of the nonprofit patent crosswalk. In section 3, I describe the nonprofit organizations that hold patents. Section 4 discusses differences the unique characteristics of nonprofit patents, employing text analysis to show that nonprofit patents are often types of technologies with positive externalities, including high spillovers, which may be under-provided by for-profit firms. Thus, nonprofit organizations may engage in research and patenting as a means of pursuing their prosocial agendas. Other possible patenting motivations are discussed in Section 5. Section 6 concludes.

2 Data

I create a crosswalk between patent identification numbers and nonprofit employer identification numbers by matching the names of patent assignees with the names of nonprofit organizations using Levenshtein Ratios of similarity. Then, I describe the types of nonprofits that hold patents and the technologies that they pertain to. Finally, I compare nonprofit patents to patents held by other types of organizations.

The nonprofit data in this paper come from the National Center for Charitable Statistics Core files, which synthesize financial and organizational data from the IRS for each year from 1989-2019. Since many organizations receive filing extensions, the data for 2018 and 2019 are incomplete. Therefore, I only use nonprofits data from 1989-2017. The Core files contain data from all organizations with more than $25,000 in gross receipts that are not religious, political, or government organizations. However, I restrict my sample to include only operating organizations, which focus on engaging in activities to support their mission, rather than distributing funds or providing private services to their members. Key variables for my analysis from this data set include organization name, employer identification number, fiscal year, royalties, and categorization by purpose from the National Taxonomy of Exempt Enties (NTEE).

The PatentsView Database from the US Patent and Trademark office provides a comprehensive data set containing information on granted patents, including the date their applications were filed, the type of patent, Cooperative Patent Classification (CPC) information, and patent number. It also contains data on the individuals and organizations that own the legal rights granted by the patent, called assignees. I consider all utility patents with assignees that were filed between 1989 and 2017, excluding patents that are assigned to publicly listed firms from the CRSP database [1], identified in Kogan et al. (2017).

I use the name standardization routine from Bessen (2020) to clean the nonprofit and patent assignee names from my sample. This process involved removing capitalization, punctuation, and accents, and standardizing common organizational descriptors. There were almost 3.4 million patent assignees and over 1.2 million unique nonprofit names.

Then, I determine how similar each patent assignee name is to each nonprofit organization name by calculating Levenshtein ratios (LR). I identified the closest match as the name with the highest LR. 85 percent of accepted matches had a LR of 100, meaning that the name of the patent assignee matched the name of a nonprofit organization perfectly. Slightly lower LRs consisted of variations of the same name (ex: BRIGHAM & WOMENS HOSPITAL INC vs. BRIGHAM & WOMENS HOSPITAL), misspelled names (ex: JOHN HOPKINS UNIVERSITY vs. JOHNS HOPKINS UNIVERSITY) and names that are similar to one another, but probably refer to different organizations (ex: WAVE ENTERPRISES INC vs. WET ENTERPRISES INC). I set acceptance thresholds in an attempt to capture observations in the first two categories while rejecting observations in the latter: I accepted LRs that were over 96, over 94 and at least 13 characters long, equal to 94 and at least 19 characters long, or equal to 93 and at least 21 characters long as matches. I also rejected 0.2 percent of the remaining observations, for which multiple organization names produced the same LR[2].

I identify 108,641 patents that belong to 2911 nonprofit organizations. There are 112,623

[1] Here, I am making the assumption that that if patent rights are shared with public for-profit firms, that patent is not counted as a nonprofit patent.

[2] This occurred because I ran my matching algorithm separately for nonprofit names that were in Core files with category code"pc" and names that were unique to the Core files with category code "pf" or "co", due to computational limitations.

observations, so about 3.5 percent of the patents are owned by multiple nonprofit organizations. Nonprofit patents account for only about 2 percent of utility patents from 1989-2017, yet technologies owned by nonprofit organizations may have an outsized impact on welfare if nonprofit patenting activity is tailored to advanced their prosocial missions.

3 Patenting Nonprofits

Nonprofit patents can be categorized according to the purpose of the organization, using the NTEE. Figure 1 shows the percent of patenting in each category. Education sector patents increased from fifty to about sixty percent of nonprofit patents, while public and societal benefit patents decreased from twenty to closer to ten percent. The health sector consistently held about twenty percent of nonprofit patents. Most education sector patents belong to institutions in higher education, and most health sector patents belong to hospitals.

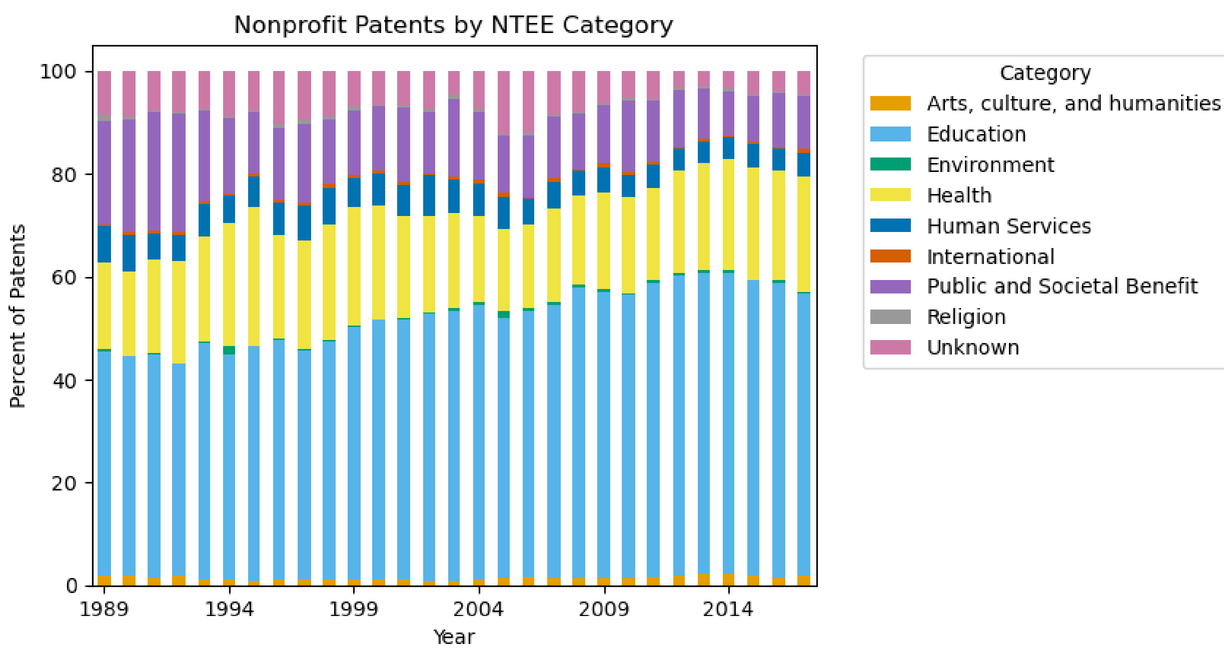

Figure 1: From the author's harmonization of data patent data from the USPTO and nonprofit data from the NCCS. NTEE stands for National Taxonomy of Exempt Entities, a categorization of nonprofit organizations according to their purpose.

While a nonprofit organization's function is related to its legal structure, organizations

Table 1: Correlation Matrix

	Nonprofit	Public For-profit	Government	Gov. Interest	Gov. Contract
Public For-profit	-0.119				
Government	0.005	-0.060			
Gov. Interest	0.355	-0.070	0.214		
Gov. Contract	0.365	-0.061	0.044	0.931	
Firm	-0.351	0.168	-0.238	-0.303	-0.266
Gov. Department	0.008	-0.070	0.704	0.164	0.028
Hospital	0.173	-0.022	-0.003	0.077	0.076
Individual	-0.005	-0.018	-0.003	-0.005	-0.005
Institute	0.205	-0.087	0.009	0.084	0.088
University	0.389	-0.132	-0.009	0.341	0.354

Notes: Assignee Characteristic correlations include data relating to the legal structure and function of the assignee organization and the patent's government involvement. Nonprofit indicates whether the name of the patent's assignee appears in a list of operating public charities from the NCCS Core Files. Patents with public for-profit assignees are identified in Kogan et al. (2017). By construction, operating nonprofit patents are not public for-profit patents, however these two categories are not exhaustive. Government represents patents with assignees in PatentsView with at least a partial interest from a US federal, county, or state government, or a foreign government. Firm, Government Department, Hospital, Individual, Nonprofit Institute, and University are functional categories suggested by key words in the organization names from Bessen (2020).

with similar functions can be for-profit. The first two columns of table 1 display the correlations between the legal categories of Nonprofit and Public For-profit[3] and the functional categories of government departments, hospitals, individuals, nonprofit institutes, universities, and firms. Since no NTEE categories are available for for-profit organizations, the functional categories represent the categorization of assignees, based on key words in their names [4]. The Hospital, Institute, and University categories are positively correlated with being nonprofit organizations and the firm category is positively correlated with being a public for-profit organization.

The USPTO classifies some assignees as government[5]. The second row of Table 1 shows that neither of the legal designations is highly correlated with being a government organization. Predictably, the functional category of Government department has a strong positive correlation with the USPTO government designation. The only other functional category correlation with the USPTO government designation that is not very small is the negative correlation with Firm.

Nonprofit organizations, especially universities, are likely to have received funding through a government contract for their patents. Government interest statements indicate which inventions were at least partially funded by a federal research grant or government contract. Having funding from government contracts is highly correlated with having a government interest statement, although the USPTO designation of an assignee as being a government is more highly correlated with having a government interest statement than with having Government contract funding. The last two columns of Table 1 show the correlations be-

[3]from Kogan et al. (2017)

[4]For example, organizations categorized as Government Departments had names that contained words like "Council" or "National"; organizations categorized as hospitals had names that contained words like "Clinic" or "Hospital"; organizations that were categorized as individuals contained words like "PhD" or "Legally Repr"; organizations that were categorized as nonprofit institutes contained words like "Foundation", "Association", or "Ministries"; organizations categorized as universities had names that contained words like "College" or "Trustees"; organizations categorized as firms had names that contained words like "& Co", "Corporation" or "International". For a complete list of keywords and the name standardization routines, see Bessen (2020)

[5]Government represents patents with assignees in PatentsView with at least a partial interest from a US federal, county, or state government, or a foreign government.

tween patents having a government interest statement or government contract funding and the functional organization categories.

4 Nonprofit Patents

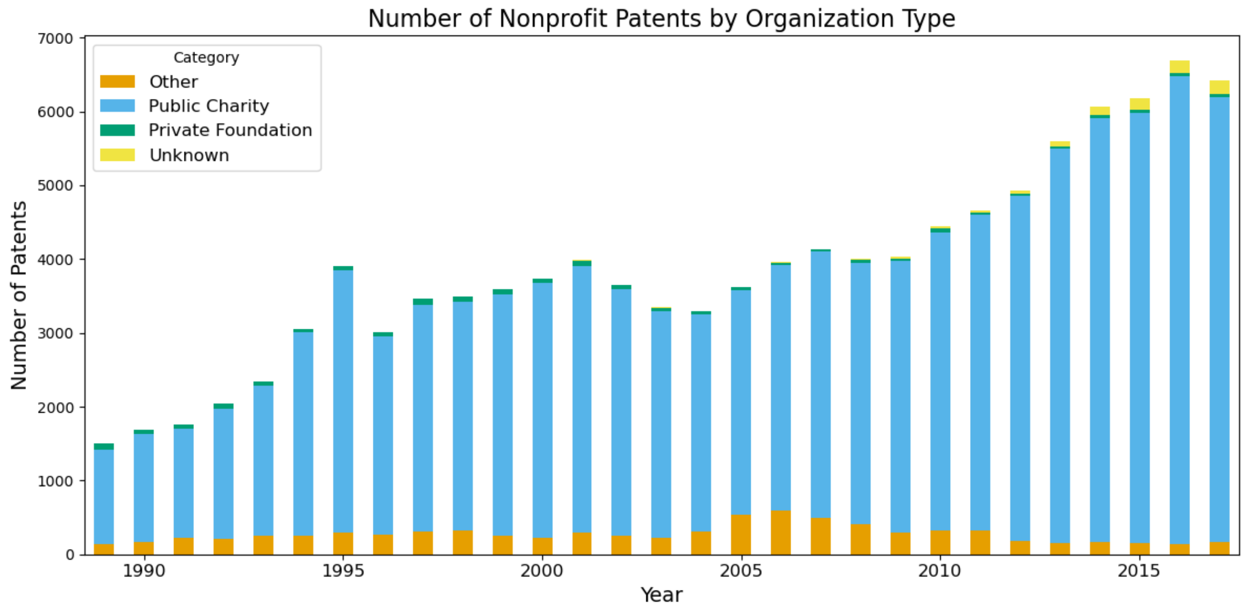

Figure 2: From the author's harmonization of data patent data from the USPTO and non-profit data from the NCCS. Public charities are organizations that receive significant public support or fall into another category, such as medical care providers, religious organizations, or educational institutions. Private foundations primarily distribute money to public charities. While Public Charities and Private Foundations are classified as 501(c)(3) organizations, other exempt organizations fall under different classifications, such as trade associations (501(c)(6)) or social and recreational clubs (501(c)(7)).

The number of nonprofit patents increased by 328 percent between 1989 and 2017, from about 1500 to 6500 patents per year. Figure 2 plots the number of patents held by nonprofit organizations that were filed each year by organization type. Most nonprofit patents are held by public charities. Bloom et al. (2020) shows that aggregate growth rates have declined, in spite of an increase in patenting beginning in the 1980's. Overall, patenting by organizations that were not nonprofits increased more slowly, by only 289 percent over the same period. Figure 3 compares the growth of patenting by nonprofits and other types of organizations. The relatively large growth of patenting by nonprofit organizations, along with the unique importance of nonprofit patents discussed in this section, suggests that the decline in innovation growth may not have been as pronounced for nonprofit organizations.

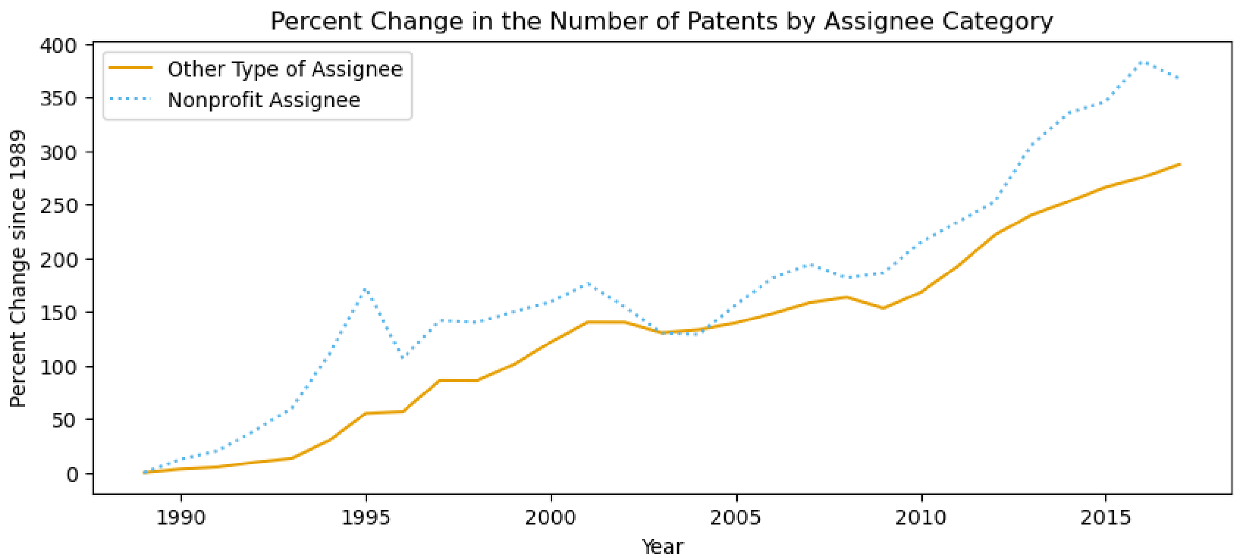

Figure 3: From the author's harmonization of data patent data from the USPTO and non-profit data from the NCCS.

4.1 CPC classes

While only two percent of patents belong to nonprofit organizations, nonprofit organizations patent disproportionately in a few CPC classes. Table 2 contains the CPC codes and descriptions of CPC codes in which nonprofit patenting is prominent. In particular, nonprofits hold about two percent of assigned patents, but over two percent of patents in these CPC classes. More than 57 percent of nonprofit patenting is in just four classes: A61, C07, G01, and C12.

Technologies may be under-provided by the market because they are public goods, have high research and development costs, or face regulatory and legal restrictions. Under-provision of technologies indicates sub-optimal levels of social welfare. By providing these technologies, nonprofit organizations may advance their own missions as well as social welfare

Table 2: CPC Class and Nonprofit Patent Participation

CPC Class	**Description**	**Nonprofit Percent of All Patents in Class**	**Percent of Nonprofit Patents in Class**
C12	Biochemistry; Beer; Spirits; Wine; Vinegar; Microbiology; Enzymology; Mutation Or Genetic Engineering	12.5	11.0
B82	Nanotechnology	12.0	0.04
C40	Combinatorial Technology	10.3	0.1
C07	Organic Chemistry	5.6	14.0
A61	Medical Or Veterinary Science; Hygiene	4.2	21.0
B81	Microstructural Technology	4.1	0.2
G16	Information And Communication Technology [ICT] Specially Adapted For Specific Application Fields	3.8	0.5
B09	Disposal Of Solid Waste; Reclamation Of Contaminated Soil	3.4	0.1
G01	Measuring; Testing	3.1	11.2
C01	Inorganic Chemistry	2.9	0.6
C30	Crystal Growth	2.8	0.2
B03	Separation Of Solid Materials Using Liquids Or Using Pneumatic Tables Or Jigs; Magnetic Or Electrostatic Separation Of Solid Materials From Solid Materials Or Fluids; Separation By High-Voltage Electric Fields	2.3	0.1
B01	Physical Or Chemical Processes Or Apparatus In General	2.3	2.2

Notes: From the author's harmonization of data patent data from the USPTO and nonprofit data from the NCCS.

by ensuring that beneficial technologies are made available.

Not only are patents in the CPC classes in Table 2 particularly likely to belong to nonprofit organizations; these CPC classes are also unusually likely to be under-provided by the market. I used the GPT-4 model from OpenAI through the API platform to determine whether each CPC class was likely to cover technologies that are public goods, have high research and development costs, or face regulatory or legal restrictions, using the prompt:

> You are a research assistant who is helping me to prepare a data set. When given a cpc class description, you will determine whether the cpc class is likely to [Question Here]. Present the final answer as a boolean object.

and the questions were as follows:

1. have high research and development costs
2. face regulatory or legal restrictions
3. cover technologies that are public goods

I provided all 128 informative CPC class titles[6] for evaluation by each prompt. Then, using the answers from these three determinations and the CPC class titles again, I asked

> You are a research assistant who is helping me to prepare a data set. When given a CPC class description and additional information, determine whether the CPC class is likely to be under-provided by the market. Consider the following: 1. Does the CPC class cover technologies that are public goods? [Response 1] 2. Does the CPC class cover technologies that have high research and development costs? [Response 2] 3. Does the CPC class cover technologies that face regulatory and legal restrictions? [Response 3] Present the final answer as a boolean object.

Table 3 presents the proportion of CPC classes that fit into each category. For thirteen CPC classes, more than two percent of patents were assigned to nonprofit organizations.

[6] There are 9 CPC class titles that are not informative, reading "Subject Matter Not Otherwise Provided in this Section" and "Technical Subjects Covered by Former USPC".

Table 3: CPC Class Underprovision Validation

	Public Good	Costly R&D	Regulations	Underprovision
	Separate Queries			*Synthesized Query*
All CPC classes	0.24	0.64	0.29	0.57
Nonprofit < 2%	0.23	0.61	0.28	0.53
Nonprofit > 2%	0.38	0.92	0.31	0.92
Top 4 Nonprofit CPC classes	0.25	1.00	0.50	1.00
F-1 Score	0.80	0.79	0.75	0.69

Notes: CPC class descriptions are provided on PatentsView. I used the GPT-4 model from OpenAI through the API platform to perform zero-shot classification of CPC class titles as having public goods, costly RD, restrictive regulations. I used these determinations in combination with the CPC class titles to determine underprovision by the market. The numbers presented in this table represent the proportion of cpc classes that are classified as fitting in the column category, for the sample indicated by the row.

These are the classes in which nonprofit patents are over represented relative to the amount of patents owned by nonprofit organizations. The second and third rows compare the subsets of CPC classes with less and more[7] than two percent of patents that were assigned to nonprofit organizations. A higher percentage of the CPC codes with more than two percent of the patents assigned to nonprofit organizations were public goods, costly to research and develop, regulated, and underprovided by markets. Almost seventy percent of all nonprofit patenting took place in only four CPC classes: Medical Or Veterinary Science; Hygiene (A61), Organic Chemistry (C07), Measuring; Testing (G01), and Biochemistry; Beer; Spirits; Wine; Vinegar; Microbiology; Enzymology; Mutation Or Genetic Engineering (C12). Row four of Table 3 shows that these four categories are even more likely to be underprovided, have high research and development costs, and face regulatory or legal restrictions.

I validated the text analysis performed by chatGPT by manually classifying each of the 128 cpc classes. Figure 4 shows the confusion matrices for the full sample and for the sample of 39 observations in which I was confident of my own determinations. Table 3 shows the F-1 scores for the confident sample.

[7]Table 2 presents the CPC classes for which more than two percent of the patents were assigned to nonprofit organizations.

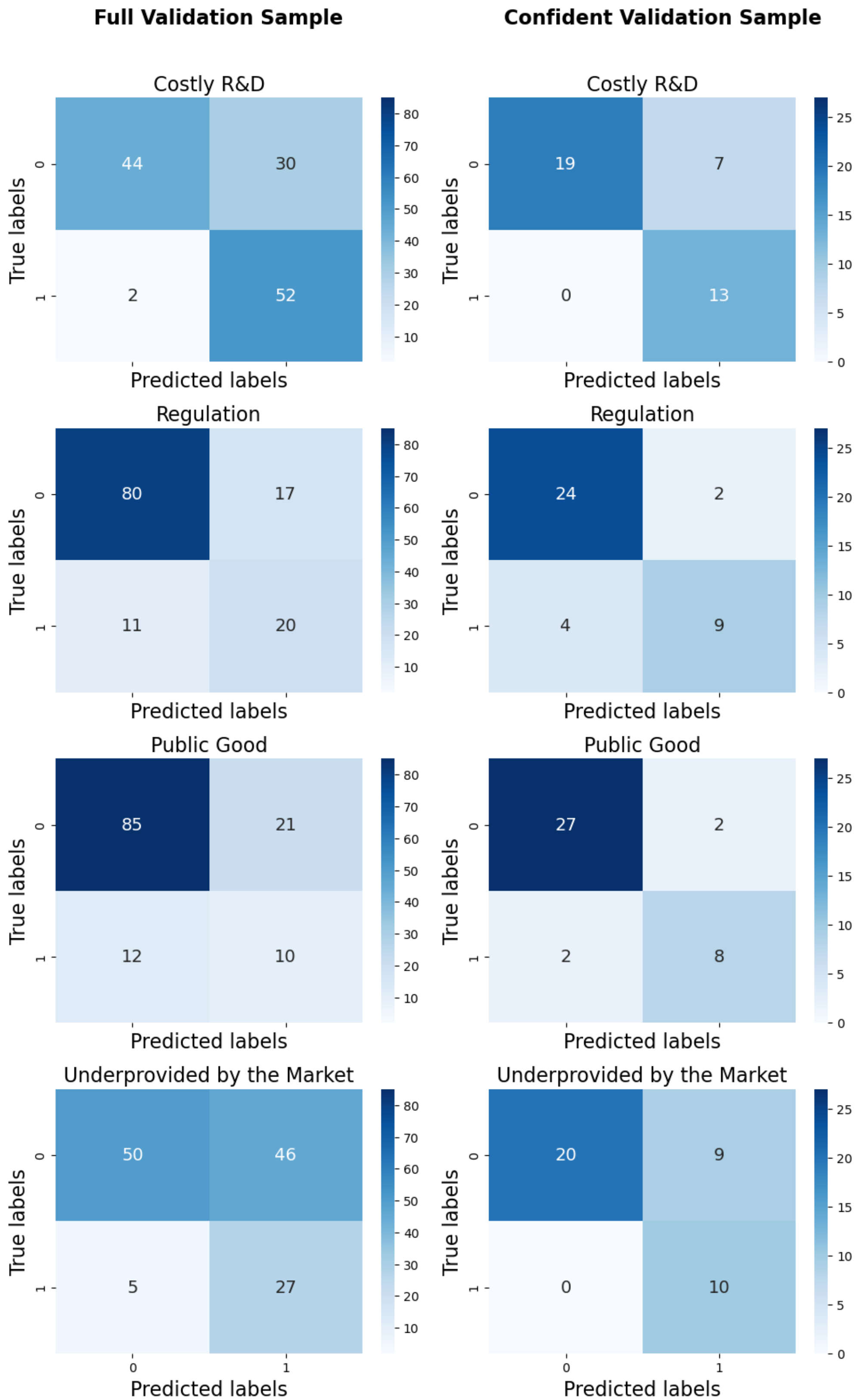

Figure 4: Confusion Matrices from Validation of text analysis of CPC Class names.

4.2 Patent Abstracts

Abstracts of nonprofit patents revealed that more nonprofit patents were green and health technologies while other organizations were more likely to hold patents related to information and communication (ICT). To determine whether patents fit into these categories, I used the GPT-4 model from OpenAI for zero-shot classification through the API platform. My prompts read as follows:

> You are a research assistant who is helping me to prepare a data set. When given a patent abstract, you will determine the answer to the question as a boolean object: [Question Here]

and the questions were as follows:

1. Is the patent for a green technology?
2. Is the patent for a healthcare technology?
3. Is the patent for a Information and communications technology?
4. Is the patent for a technology that will facilitate future research?

In addition to prominent categories of technology, I also used GPT-4 to determine how basic the research described in the abstract was, using the following prompt:

> You are a research assistant. When given a patent abstract, you will determine whether it describes pure applied research or use-inspired basic research. Pure applied research belongs in Edison's quadrant, and use-inspired basic research belongs in Pasteur's quadrant. The patent must belong to exactly one of these two categories. Give the answer True if the patent is pure applied research and False if the patent is use-inspired basic research.

Pasteur's quadrant and Edison's quadrant refer to a canonical classification of scientific research along axies of usefulness and basicness from Stokes (1997). Only Edison's quadrant and Pasteur's quadrant are applicable because patents are only granted for useful inventions.

I then provided a randomly selected sample of 1000 nonprofit and 1000 other patent abstracts, individually, for evaluation by each prompt. Table 4 shows my results. The "other" category contains public for-profit patents. Results for this exercise on the public for-profit subset of the other organizations can be found in Table 11 in the appendix.

As noted in Table 4, nonprofit patents are much more likely to be in the green or health fields than other patents. Health technologies are under provided by the market because positive externalities, reservation of inventions for emergency situations, and lengthy trials prevent innovators from capturing the full value of their innovations during the patent protection period (J.B. Rebitzer and R.S. Rebitzer, 2023). Provision of health technologies is particularly consistent with the missions of organizations in the health NTEE category. Green technologies are also likely to be under-provided because without a proper carbon price, dirty technology captures an inappropriately high market share (Acemoglu et al., 2012). Also over represented in nonprofit patents, fundamental science may be under-provided by the private sector if it is very costly to produce or if the results and commercial applications are uncertain (Akcigit, Hanley, and Serrano-Velarde, 2021).

About eight percent of nonprofit patents would be considered use-inspired basic research, while less than one percent of other patents would. Akcigit, Hanley, and Serrano-Velarde (2021) determine that basic research has particularly large spillovers. Based on the structure of my inquiry, Pure Applied research is simply the inverse of use-inspired basic research. Furthermore, about 49 percent of nonprofit patents were for a technology that will facilitate future research, while only 16 percent of other patents would facilitate future research. Hence, this analysis of patent abstracts provides evidence that nonprofit patents are also concentrated in high-spillover technologies.

To validate the output from chatGPT, I drew a random sample of 200 from the 2000 patents that I asked chatGPT to classify. My random sample included 102 nonprofit patents and 98 patents that belonged to other types of organizations. Then, I manually determined whether a patent is a Green, Medical, or Information and Communication Technol-

Table 4: Comparison between Patent Abstracts pertaining to Technology Categories and Invention Characteristics

	Nonprofit	Other	Difference	F-1 Score
% Green	6.3	2.9	3.4	1.00
% Health	49.1	11.7	37.4	0.96
% ICT	15.1	37.5	-22.4	0.98
% Future Research	48.5	15.5	33.0	0.76
% Pure Applied	92.4	99.4	-7.0	0.94
% Use-inspired Basic	7.6	0.6	7	0.33

Notes: Nonprofit patents were determined using the author's harmonization of data patent data from the USPTO and nonprofit data from the NCCS. A sample of 1000 nonprofit and 1000 not nonprofit abstracts were analyzed using the GPT-4 model from OpenAI through the API platform to perform zero-shot classification. Testing the difference between the nonprofit and other samples indicated that the differences between nonprofit patents and patents belonging to other types of organizations were significant at the 1% level in all cases.

ogy, whether it facilitates future research, or represents pure applied research. I eliminated potential bias by ensuring that I knew neither whether the patent had a nonprofit assignee nor chatGPT's categorization. I also created a sub-sample of 71 patents for which I was confident of my categorization. Then, I compared my validation responses to the responses of chatGPT for both the random sample and the sub sample of patents of which I was confident. Confusion matrices from my validation are displayed in Figure 5. The right column of confusion matrices in figure 5 is from the full validation sample, and the left column is from the sample of abstracts that I classified with confidence. The last column of Table 4 displays the F-1 scores for each category for the confident validation sample. I consider F-1 scores above 0.7 good, especially given the abstract nature of these determinations.

4.3 Citations

The dissemination of patented ideas can be measured by counting the number of citations that they receive. The regressions in Table 5 show that nonprofit patents have more citations on average than other patents. I follow several authors in controlling for the time since the patent was filed and the sector to which the technology and assignee belong (Hall, A.B. Jaffe, and Trajtenberg, 2001; Akcigit and Kerr, 2018; Lerner and Seru, 2022). Comparing Mod-

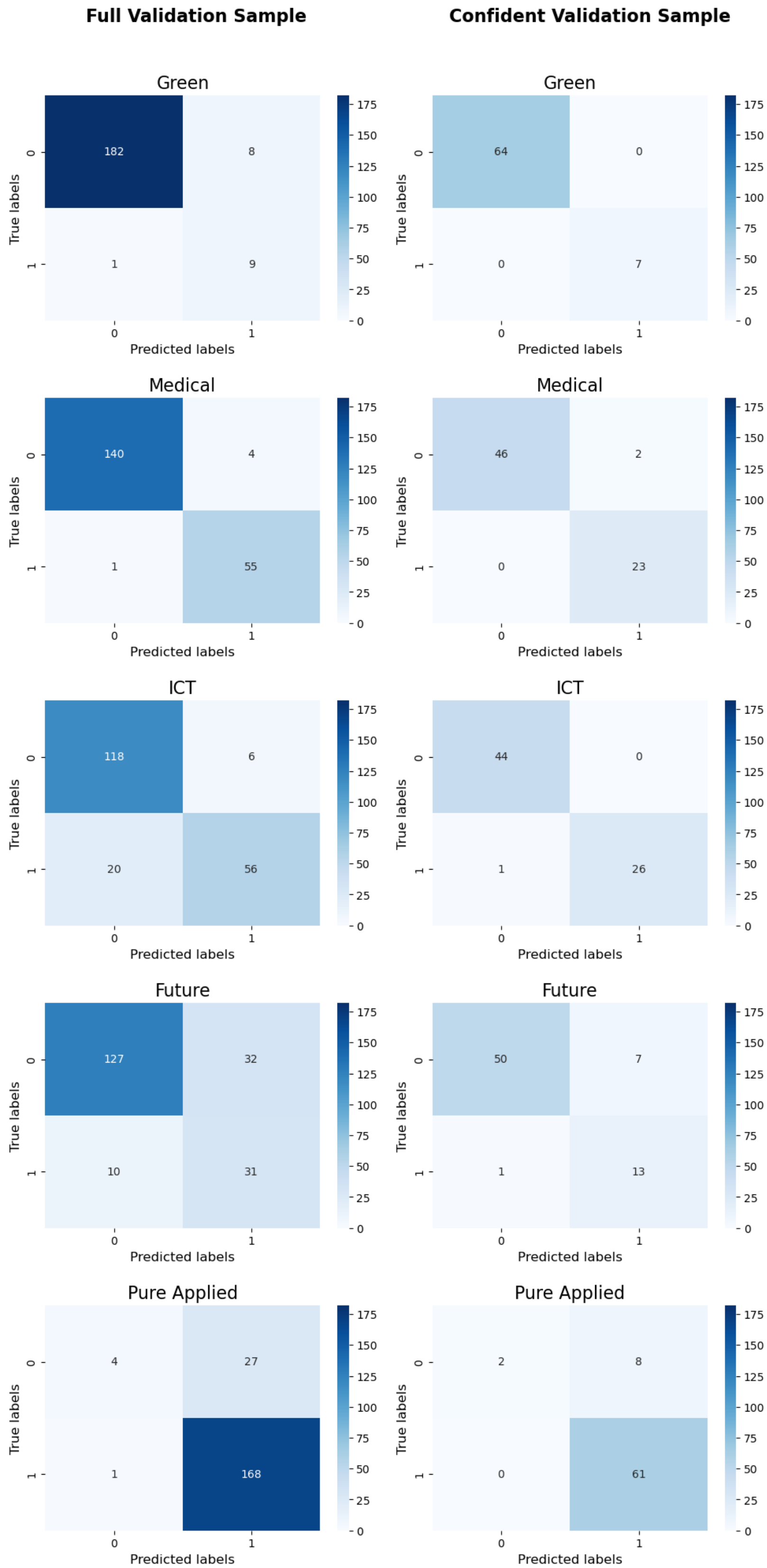

Figure 5: Confusion Matrices from Validation of text analysis of Patent Abstracts for Non-profit and Other types of organizations.

Table 5: Citations Regression Results

	Dependent variable: Number of Citations				
	Model 1	Model 2	Model 3	Model 4	Model 5
Model 6					
Nonprofit	4.670***	1.576***	2.733***	3.687***	3.253***
	(0.193)	(0.192)	(0.193)	(0.219)	(0.223)
Public For-profit			3.133***	3.047***	3.038***
			(0.059)	(0.060)	(0.060)
Government			-5.559***	-3.054***	-3.862***
			(0.332)	(0.491)	(0.497)
Name(Gov)				-2.867***	-2.995***
				(0.406)	(0.406)
Name(Hosp)				-5.964***	-6.314***
				(0.770)	(0.771)
Name(Indiv)				-6.948***	-6.932***
				(0.860)	(0.860)
Name(Inst)				-4.049***	-4.167***
				(0.241)	(0.241)
Name(Univ)				-0.103	-0.650***
				(0.193)	(0.200)
Government Interest					2.130***
					(0.202)
CPC class		✓	✓	✓	✓
File Year	✓	✓	✓	✓	✓
const	27.163***	24.003***	23.184***	23.265***	23.296***
	(0.194)	(0.337)	(0.337)	(0.337)	(0.337)
Observations	3652733	3652733	3652733	3652733	3652733
R^2	0.043	0.084	0.085	0.085	0.085
Adjusted R^2	0.043	0.084	0.085	0.085	0.085

Notes: Results from regressions of the number of citations received by a patent on whether the patent was an operating nonprofit organization and other explanatory variable. The number of citations, filing year, CPC class (a system of classifying types of technology), Government assignee, and government interests are based on data from PatentsView. Nonprofit indicates whether the name of the patent's assignee appears in a list of operating public charities from the NCCS Core Files. Patents with public for-profit assignees are identified in Kogan et al. (2017). By construction, operating nonprofit patents are not public for-profit patents. Patents that are assigned to operating nonprofits or public for-profit organizations have more citations than patents with assignees in neither category. Patents with government assignees had fewer citations. The name standardization routines in Bessen (2020) categorize patent assignees by organization type based on key words in their names. The categories are government departments, hospitals, individuals, nonprofit institutes, universities, and firms. These categories suggest the function of each organization but are not directly related to the legal designation of the organization as a nonprofit. The negative coefficients on the categories other than firms implies that, all else equal, their patents receive fewer citations than patent assigned to firms. *p<0.1; **p<0.05; ***p<0.01

els 1 and 2 reinforces that one way nonprofit organizations advance their missions through patenting is by specializing in categories of technology that, in addition to potentially being under-provided by the market, have high levels of citations. This could be either because these technologies particularly benefit from spillovers or because nonprofit involvement encourages the dissemination of ideas. Even when controlling for CPC class, nonprofit patents receive more citations than other patents.

The positive relationship between being assigned to a nonprofit organization and citations increases when I control for more organization characteristics. Similar to nonprofit patents, patents with public for-profit assignees have more citations than patents with assignees in neither category in Model 3. In contrast, patents that were assigned to government organizations had fewer citations[8]. In Model 4, the negative coefficients on the functional categories[9] based on organization names implies that, all else equal, patents from governments, hospitals, nonprofit institutes, individuals, and universities receive fewer citations than patents assigned to firms. When controlling for the assignee's function, patents owned by nonprofits still receive more citations than other types of organizations, including public for-profits.

Without government involvement, technology some types of technology may be under provided by the market (Bernanke, 2011). Moreover, the 1980 Bayh-Dole act ensures that government-funded inventions are used in practical applications, which could affect the transmission and citations of their patents (TRAIN, 2020). Indeed, government interest statements have a positive effect on citations in Models 5 of Table 5. Comparing Model 5 to Model 4 in Table 5, the main effect of adding government interest is that the coefficient on university becomes more significantly negative. Adding government interest has little effect on the positive relationship between between a patent's belonging to a nonprofit organization

[8]In table 5, classification as a government organization does not exclude a patent from being assigned to a nonprofit or public for profit organization. Less than one percent of nonprofit or public for profit patents were also considered government patents. As robustness, to ensure that the small overlap between government and other organizational legal types does not drive results, Table 10 in the Appendix re-runs the regressions in Table 5, removing nonprofit and public for profit designations from any patents that also had a government assignee. The results are very similar.

[9]See the discussion of Table 1 for a description of functional categories.

and number of citations.

Table 7 in the Appendix presents the same regressions as Table 5, but with additional interaction terms between the functions indicated by organization names and the legal designations of Nonprofit, Public For-profit, and Government organizations. Adding the interaction terms has very little effect on the coefficients and no effect on the significance levels of the coefficients on the variables in Table 5. The only type of nonprofit institution that has fewer citations, when taking into account the interaction term, are government organizations.

Excluding self-citations may improve the measurement of idea dissemination by patents. When I replace the dependent variable in my regressions with the number of citations by patents assigned to other organizations, as shown in Table 8 in the Appendix, the coefficients on the types of legal organizations remain similar in direction and significance, but decrease in magnitude. Rather than having 3.8 additional citations compared to patents belonging to firms with unidentified organizational legal structure, nonprofit patents have 2.4 additional out-of-organization citations in Model 5. As in Table 7, only patents assigned to nonprofit organizations classified as government, based on their names, had fewer citations than patents assigned to the organizations of the same function but no data on the legal category.

Number of second-generation citations is another way to measure the dissemination of ideas by patents. Appendix Table 9 contains the results of regressions with second-generation citations as the dependent variable. The first three models show that nonprofit organizations receive fewer citations when accounting for legal structure and not organization function. However, Models 4 and 5 reveal that this does not apply to nonprofit Universities and Hospitals. Nonprofit hospitals, in particular, have higher second generation citations. This is consistent with a prosocial response from nonprofit hospitals to a push in the health industry for organizations with patents to adopt open or non-exclusive licenses, which is happening to ensure access to critical technologies in developing countries and public health crises ('t Hoen, 2009) and to facilitate rapid development across many contributors (Ledford, 2013 and Henry et al., 2002).

Table 6: Patent Importance

	Backward Similarity (5 yrs)	Forward Similarity (10 yrs)	Importance
Nonprofit	4.472e+06	7.713e+06	2.093
Other	6.147e+06	1.015e+07	1.871

Notes: Backward Similarity and Forward Similarity are measurements of patent novelty and impact based on the similarity between the text of patent documents from Kelly et al. (2021). Backward Backward Similarity (5 yrs) covers all patents filed in the five years prior to that patent's filing, and Forward Similarity (10 yrs) covers all patents filed within 10 years of a patent's filing. Patent Importance is measured as the ratio between Backward Similarity (5 yrs) and Forward Similarity (10 yrs). The numbers in this table are the means over patents from 1989-2010 according to the author's categorization of the patent's assignee being a nonprofit organization.

4.4 Importance

In Kelly et al. (2021), important patents are "distinct improvements in the technological frontier [that] become the new foundation upon which subsequent inventions are built". Relating this concept to high spillover technologies, I find that nonprofit patents have higher importance than patents assigned to other types of organizations in Table 6. The patent importance indicator is the ratio between a patent's impact and novelty. Impact is measured by the similarity between a patent's text and the text of patents filed in the ten years after it. Nonprofit patents had lower forward similarity than patents with other assignees, so they had lower impact. However, a patent must be both impactful and novel to be important. Novelty is measured by the similarity between a patent's text and the text of patents filed in the five years prior to it. Nonprofit patents had lower backwards similarity than patents with other assignees, so they were more novel, and also more important. Indeed, they make up over eight percent of all breakthrough patents identified by Kelly et al. (2021) between 1989 and 2002, while making up only less than 2.5 percent of all patents over the same period.

5 Why Nonprofits Patent

Patents that are assigned to nonprofit organizations are particularly welfare improving because they tend to be in under-provided technology categories and have positive externalities, including high spillovers. Hence, developing and patenting these technologies may be consistent with a nonprofit organization's prosocial mission. Nevertheless, it is possible that other motivations could also drive nonprofit innovation and patenting. Although nonprofits explicitly pursue social objectives, rather than profit, nonprofit organizations have a profit motive analogue: the royalties can fund other programs that pursue these objectives. Patents can also provide reputation benefits to inventors, even if they are owned by the inventor's employer. Defensive patenting can protect the nonprofit's ability to use its own innovations in pursuit of its objective.

Analysis of nonprofit royalties suggests that the profit motive analogue does not drive nonprofit patenting. Alon, Capelle, and Matsuda (2023) find that between 1991 and 2018, no university earned more than 40 percent of its research expenditures in licensing revenue, and the median university earned less than two percent. Hence, revenue from patent royalties could not be motivating university research spending. Nonprofits with patents, in the education NTEE category, collected about twice as much royalties income per organization as organizations in any other NTEE category between 2011 and 2017. Thus, as in education, patent-producing research is most likely a use, rather than a source, of funds in the rest of the nonprofit sector.

While the provision of socially beneficial technologies seems to motivate nonprofit research, the specific act of patenting these technologies may be driven by reputation benefits for the inventors, or be defensive for the organization. Even if an organization does not collect revenue from licensing their patents [10] or exclude competitors, defensive patenting protects an organization's ability to use its own technologies without paying licensing fees or fighting legal battles over patent infringement with assignees of related patents (Schultz and Urban, 2012). Costs associated with patenting driven by inventor incentives can be considered part of the cost of hiring researchers to develop prosocial technologies. Neither of these patenting motives detracts from the nonprofit's goal of developing and disseminating

[10] According to my data, this is the case from 2011-2017 for almost 90 percent of nonprofits that patented between 1989 and 2017.

welfare-improving technologies.

Research activities can also provide auxiliary benefits towards a nonprofit organization's mission. For example, Alon, Capelle, and Matsuda (2023) points out that research and patenting can help improve the quality of education that a nonprofit university provides. Such auxiliary benefits of patenting do not detract from nonprofit innovators' position as a leader in welfare-improving innovations.

6 Conclusion

This paper is the first to describe nonprofit patenting. I identify patents that are assigned to nonprofit organizations by matching assignee names to the names of nonprofits that filed 990s between 1989 and 2017. I find that education (58%), health (22%), and public and societal benefit organizations (15%) make up the majority of nonprofit organizations that are involved in patenting. About seventy percent of nonprofit patents are in just four CPC classes, which relate to basic scientific and health research.

The existence and acceleration of nonprofit patenting requires greater nuance in economic theories regarding patenting. Nonprofit patenting has expanded more rapidly than patenting by other entities, indicating potential disparities in innovation growth rates across different types of organizations. Furthermore, the profit motive may not be sufficient for explaining all patenting, as it does in endogenous growth theory.

Data from nonprofit patenting suggest that technologies invented by nonprofits are in under-provided classes and have higher externalities. Specifically, nonprofit organizations patent disproportionately in health and green technologies, both of which are not provided by the market at a socially optimal level. Nonprofit patents also tend to be closer to basic research and receive more citations, both of which suggest strong spillovers from nonprofit technologies. These characteristics suggest that nonprofit technologies may make greater-than-average welfare contributions, consistent with their prosocial missions. Hence, endogenous growth models and the policies that they inform should take nonprofit motives and incentives into account.

References

't Hoen, E.F.M. (2009) *The Global Politics of Pharmaceutical Monopoly Power*. AMB Publishers.

Acemoglu, D. (2023) "Distorted Innovation: Does the Market Get the Direction of Technology Right?" *AEA Papers and Proceedings*. Vol. 113, pp. 1–28.

Acemoglu, D. et al. (2012) The environment and directed technical change. *American economic review* 102.1, 131–166.

Akcigit, U., D. Hanley & N. Serrano-Velarde (2021) Back to basics: Basic research spillovers, innovation policy, and growth. *The Review of Economic Studies* 88.1, 1–43.

Akcigit, U. & W.R. Kerr (2018) Growth through heterogeneous innovations. *Journal of Political Economy* 126.4, 1374–1443.

Alon, T., D. Capelle & K. Matsuda (2023) University Research and the Market for Higher Education. *Working Paper.*

Autor, D. et al. (2020) Foreign competition and domestic innovation: Evidence from US patents. *American Economic Review: Insights* 2.3, 357–374.

Bernanke, B.S. (2011) Promoting Research and Development The Government's Role. *Issues in Science and Technology* 27.4, 37–41.

Bessen, J. (2020) *Name Standardization Routines Uploaded*. NBER Patent Data Project.

Bloom, N. et al. (2020) Are ideas getting harder to find? *American Economic Review* 110.4, 1104–1144.

Core Files (2021). (accessed May 23, 2022). Washington, DC: Urban Institute, National Center for Charitable Statistics.

Donald, E. (2023) Spillovers and the Direction of Innovation: An Application to the Clean Energy Transition.

Hall, B.H., A.B. Jaffe & M. Trajtenberg (2001) *The NBER patent citation data file: Lessons, insights and methodological tools.*

Henry, M.R. et al. (2002) DNA Patenting and Licensing. *Science* 297.5585, 1279–1279.

Kalyani, A. (2022) The Creativity Decline: Evidence from US Patents. *Available at SSRN 4318158.*

Kelly, B. et al. (2021) Measuring technological innovation over the long run. *American Economic Review: Insights* 3.3, 303–320.

Kogan, L. et al. (2017) Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics* 132.2, 665–712.

Krieger, J.L., M. Schnitzer & M. Watzinger (2022) *Standing on the Shoulders of Science.* Tech. rep. Harvard Business School Working Paper.

Ledford, H. (2013) Bioengineers Look Beyond Patents. *Nature* 499, 16–17.

Lerner, J. & A. Seru (2022) The use and misuse of patent data: Issues for finance and beyond. *The Review of Financial Studies* 35.6, 2667–2704.

Rebitzer, J.B. & R.S. Rebitzer (2023) *Why Not Better and Cheaper?: Healthcare and Innovation.* Oxford University Press.

Romer, P.M. (1990) Endogenous technological change. *Journal of political Economy* 98.5, Part 2, S71–S102.

Schultz, J. & J.M. Urban (2012) Protecting open innovation: The defensive patent license as a new approach to patent threats, transaction costs, and tactical disarmament. *Harv. JL & Tech.* 26, 1.

Stokes, D. (1997) Pasteur's Quadrant: Basic Science and Technological Innovation (The Brookings Institution, Washington DC).

The Research Acceleration and Innovation Network (TRAIN) (2020) *The Bayh-Dole Act of 1980: Frequently Asked Questions.* The Milken Institute.

Trajtenberg, M., R. Henderson & A. Jaffe (1997) University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and new technology* 5.1, 19–50.

USPTO (2023) *PatentsView*.

Appendix

Citations Regressions

Table 7: Citations Regression Results With Interactions

	Dependent variable: Number of Citations per Patent				
	Model 1	Model 2	Model 3	Model 4	Model 5
Nonprofit	4.670***	1.576***	2.733***	4.052***	3.779***
	(0.193)	(0.192)	(0.193)	(0.352)	(0.353)
Public For-profit			3.133***	3.044***	3.037***
			(0.059)	(0.060)	(0.060)
Government			-5.559***	-6.395***	-6.924***
			(0.332)	(1.215)	(1.216)
Name(Gov)				-2.825***	-2.902***
				(0.463)	(0.463)
Name(Hosp)				-9.802***	-9.910***
				(1.823)	(1.823)
Name(Indiv)				-7.365***	-7.346***
				(0.908)	(0.908)
Name(Inst)				-4.326***	-4.330***
				(0.289)	(0.289)
Name(Univ)				0.066	-0.485**
				(0.227)	(0.233)
Nonprofit*Name(Gov)				-4.157***	-3.797**
				(1.597)	(1.598)
Nonprofit*Name(Hosp)				2.952	2.508
				(2.060)	(2.060)
Nonprofit*Name(Indiv)				10.105	9.946
				(23.733)	(23.733)
Nonprofit*Name(Inst)				0.753	0.232
				(0.622)	(0.624)
Nonprofit*Name(Univ)				-0.929*	-1.072**
				(0.498)	(0.498)
Government Interest					2.097***
					(0.203)
For-Profit*Names				✓	✓
Government*Names				✓	✓
CPC class		✓	✓	✓	✓
File Year	✓	✓	✓	✓	✓
Constant	27.163***	24.003***	23.184***	23.269***	23.295***
	(0.194)	(0.337)	(0.337)	(0.337)	(0.337)
Observations	3652733	3652733	3652733	3652733	3652733
Adjusted R^2	0.043	0.084	0.085	0.085	0.085

Notes: Results from regressions of the number of citations received by a patent on weather the patent was an operating nonprofit organization and other explanatory variables. The explanatory variables included in this table are the same as in Table 5, but the regressions include interactions between the functions indicated by organization names and the legal designations of nonprofit, Public For-profit, and Government organizations. *p<0.1; **p<0.05 ; ***p<0.01

Table 8: Citations Excluding Self-cites Regression Results

	Dependent variable:				
	Number of Citations per Patent, Excluding Self-cites				
	Model 1	Model 2	Model 3	Model 4	Model 5
Nonprofit	2.348***	0.664***	1.397***	2.609***	2.410***
	(0.114)	(0.115)	(0.115)	(0.218)	(0.219)
Public For-profit			2.254***	2.175***	2.171***
			(0.037)	(0.037)	(0.037)
Government			-3.937***	-5.086***	-5.447***
			(0.206)	(0.760)	(0.761)
Name(Gov)				-1.887***	-1.922***
				(0.259)	(0.259)
Name(Hosp)				-6.792***	-6.834***
				(0.922)	(0.922)
Name(Indiv)				-5.061***	-5.050***
				(0.565)	(0.565)
Name(Inst)				-2.921***	-2.923***
				(0.161)	(0.161)
Name(Univ)				-0.498***	-0.771***
				(0.120)	(0.123)
Nonprofit**Name*(Gov)				-2.779***	-2.583***
				(0.911)	(0.911)
Nonprofit**Name*(Hosp)				2.006*	1.704
				(1.078)	(1.078)
Nonprofit**Name*(Indiv)				11.085	11.019
				(17.565)	(17.564)
Nonprofit**Name*(Inst)				0.273	-0.056
				(0.373)	(0.374)
Nonprofit**Name*(Univ)				-0.705**	-0.851***
				(0.294)	(0.295)
Government Interest					1.291***
					(0.122)
For-Profit*Names				✓	✓
Government*Names				✓	✓
CPC class		✓	✓	✓	✓
File Year	✓	✓	✓	✓	✓
Constant	23.543***	21.858***	21.247***	21.317***	21.328***
	(0.140)	(0.206)	(0.207)	(0.207)	(0.207)
Observations	5412417	5412417	5412417	5412417	5412417
Adjusted R^2	0.072	0.100	0.101	0.101	0.101

Notes: Results from regressions of the number of citations by patents assigned to other organizations received by a patent on whether the patent was an operating nonprofit organization and other explanatory variables. The explanatory variables included in this table are the same as in Table 7. *p<0.1; **p<0.05 ; ***p<0.01

Table 9: Second-generation Citations Regression Results

	Dependent variable: *Number of Second-generation Citations per Patent*				
	Model 1	Model 2	Model 3	Model 4	Model 5
Nonprofit	-244.831***	-693.524***	-469.931***	-258.915***	-260.581***
	(45.766)	(46.374)	(46.676)	(85.356)	(85.596)
Public For-profit			596.290***	580.718***	580.676***
			(14.325)	(14.519)	(14.520)
Government			-23.931	27.010	23.787
			(80.478)	(294.225)	(294.486)
Name(Gov)				-145.805	-146.274
				(112.094)	(112.109)
Name(Hosp)				-1711.127***	-1711.788***
				(441.649)	(441.657)
Name(Indiv)				-164.521	-164.410
				(219.815)	(219.816)
Name(Inst)				-187.833***	-187.859***
				(69.922)	(69.922)
Name(Univ)				-619.783***	-623.140***
				(55.021)	(56.516)
Nonprofit* Name(Gov)				-66.013	-63.818
				(386.855)	(386.947)
Nonprofit* Name(Hosp)				753.868	751.164
				(498.858)	(498.966)
Nonprofit* Name(Indiv)				-104.996	-105.967
				(5748.428)	(5748.430)
Nonprofit* Name(Inst)				-64.896	-68.071
				(150.599)	(151.093)
Nonprofit* Name(Univ)				278.351**	277.479**
				(120.514)	(120.561)
Government Interest					12.780
					(49.160)
For-Profit* Names				✓	✓
Government* Names				✓	✓
CPC class		✓	✓	✓	✓
File Year	✓	✓	✓	✓	✓
Constant	48.164	-329.211***	-506.085***	-489.058***	-488.895***
	(46.127)	(81.599)	(81.713)	(81.731)	(81.734)
Observations	3652733	3652733	3652733	3652733	3652733
Adjusted R^2	0.004	0.009	0.009	0.009	0.009

Notes: Results from regressions of the number of second-generation citations received by a patent

Table 10: Citations Regression Results, Government Exclusive

	Dependent variable: Number of Citations per Patent				
	Model 1	Model 2	Model 3	Model 4	Model 5
Nonprofit	4.743***	1.630***	2.739***	3.707***	3.267***
	(0.194)	(0.192)	(0.193)	(0.220)	(0.224)
Public For-Profit			3.132***	3.045***	3.037***
			(0.059)	(0.060)	(0.060)
Government			-5.395***	-2.837***	-3.656***
			(0.332)	(0.491)	(0.497)
Name(Gov)				-2.902***	-3.027***
				(0.406)	(0.406)
Name(Hosp)				-5.979***	-6.324***
				(0.770)	(0.771)
Name(Indiv)				-6.949***	-6.933***
				(0.860)	(0.860)
Name(Inst)				-4.051***	-4.168***
				(0.241)	(0.241)
Name(Univ)				-0.112	-0.655***
				(0.193)	(0.200)
Government Interest					2.125***
					(0.202)
CPC Class		✓	✓	✓	✓
File Year	✓	✓	✓	✓	✓
const	27.162***	24.003***	23.184***	23.266***	23.296***
	(0.194)	(0.337)	(0.337)	(0.337)	(0.337)
Observations	3652733	3652733	3652733	3652733	3652733
R^2	0.043	0.084	0.085	0.085	0.085
Adjusted R^2	0.043	0.084	0.085	0.085	0.085

Notes: Results from regressions of the number of citations received by a patent on weather the patent was an operating nonprofit organization and other explanatory variables. The explanatory variables included in this table are the same as in Table 5, but the regressions include interactions between the functions indicated by organization names and the legal designations of nonprofit, Public For-profit, and Government organizations. For this robustness check on Table 5, the legal categories Nonprofit, Public For-profit, Government, and Private For-profit were made to be mutually exclusive by excluding Government patents from also being Nonprofit patents. *p<0.1; **p<0.05 ; ***p<0.01

Nonprofit vs For-profit Patent Abstracts

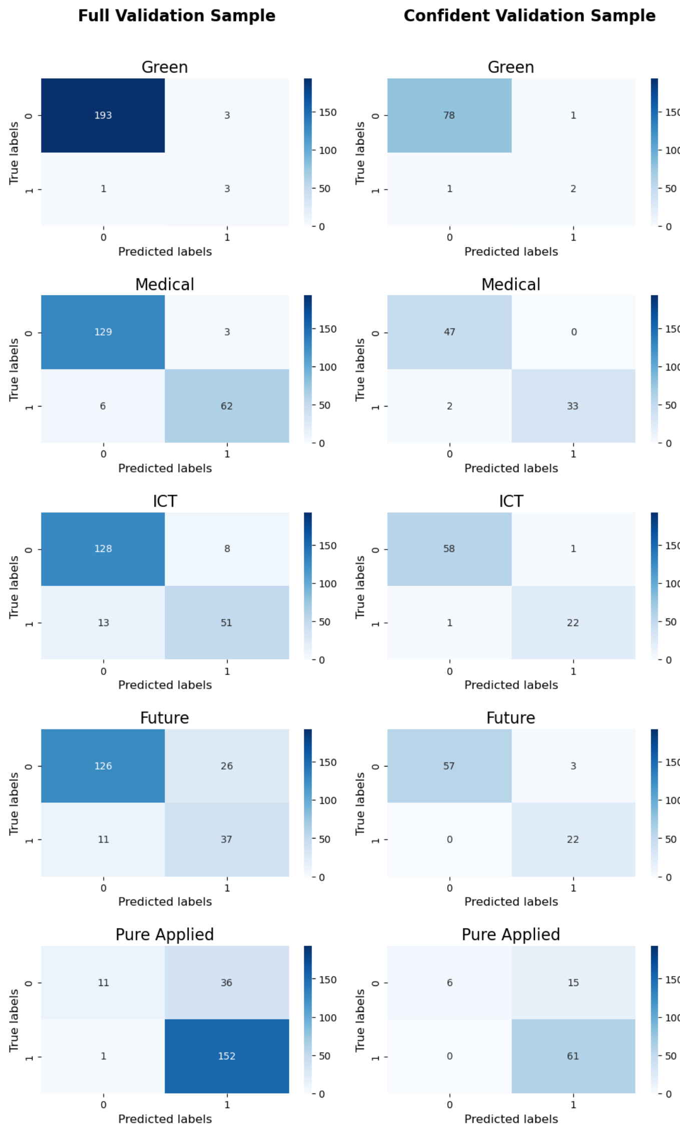

Figure 6: Confusion Matrices from Validation of text analysis of Patent Abstracts for Non-profit and Public For-Profit organizations.

Table 11: Comparison between Patent Abstracts pertaining to Technology Categories and Invention Characteristics

	Nonprofit	**Public For-Profit**	**Difference**	**F-1 Score**
% Green	6.3	2.5	3.8	0.67
% Health	49.1	7.7	41.4	0.97
% ICT	15.1	47.8	-32.7	0.96
% Future Research	48.5	17.1	31.4	0.94
% Pure Applied	92.4	99.2	-6.8	0.89
% Use-inspired Basic	7.6	0.8	6.8	0.44

Notes: Nonprofit patents were determined using the author's harmonization of data patent data from the USPTO and nonprofit data from the NCCS. A sample of 1000 nonprofit and 1000 public for profit patent abstracts were analyzed using the GPT-4 model from OpenAI through the API platform to perform zero-shot classification. Testing the difference between the nonprofit and Public For-Profit samples indicated that the differences between for-profit and nonprofit patents were significant at the 1% level in all cases.