Università degli Studi di Torino

Innovation for the Circular Economy Doctoral Program

XXXIV Cycle

Essays on patent analytics

by

Grazia Sveva Ascione

Supervisor:

Professor Francesco Quatraro

Department of Economics and Statistics "Cognetti de Martiis"

University of Turin

Essays on patent analytics

by

Grazia Sveva Ascione

PhD in Innovation for the Circular Economy

Introduction

This thesis examines and combines the recent rise of two phenomena, patent litigation and the transition to green and sustainable technologies. I examine these phenomena from an economic and data science perspective using United States (US) patents with their numerical and textual attributes. The chapters of this thesis highlight the evolution of the litigation phenomenon when multiple actors are involved and the interplay with the rapidly evolving landscape of green and sustainable technologies, concluding with an in-depth analysis of technologies in the context of the Sustainable Development Goals (SDGs).

The first chapter-co-authored with Laura Ciucci, Claudio Detotto, and Valerio Sterzi-examines how, in recent years, some universities have begun to pursue "overzealous" strategies to protect their existing patents by enforcing them in court and selling them to the highest bidder. In this chapter, we provide the first comprehensive evidence on the characteristics of universities' litigation strategies by comparing patents litigated by universities with those litigated by patent trolls and other entities. To do so, we collect data on patent infringement lawsuits in the United States from 2003-2016 and analyze three dimensions that have been identified in the literature as characteristics of patent trolls' behavior: (i) the intensity with which a patent is litigated, (ii) the decision to file a patent suit in the Federal District Court of Texas Eastern, and (iii) the quality of the patents asserted. Our results show that while universities' overall litigation strategies appear to differ from those of patent trolls, this is not the case in the information and communication technologies sector, which is most targeted by trolls, as universities often file their patents in the Eastern District of Texas and these are of lower quality compared to patents filed by other institutions.

The second chapter - co-authored with Francesco Quatraro and Valerio Sterzi - empirically examines the impact of green public procurement (GPP) on patent litigation dynamics in the US. Green technologies have gained prominence over the past twenty years due to increased policy attention to the ecological transition. At the same time, patent infringement lawsuits in the US have increased at an alarming rate since 2009, shedding light on their potential side effects on the innovation ecosystem. In this research, we use detailed geographic data sources on green patents, procurement expenditures at the US county level, and litigation data for the period 2003-2016. We reach two conclusions: first, we find that green patents are not at high risk of litigation. This could be due to the inherent complexity of green patents, which has been associated with a lower likelihood of litigation. Second, when we look at litigation involving patent assertion entities (PAE), higher GPP is associated with a higher likelihood of litigation. This suggests that while stricter environmental regulations are intended to encourage the adoption and market for green technologies, they actually favor strategic behaviors by those who own these technologies. The policy implications of these findings call for more careful consideration of these "side effects" of stricter environmental regulations, which in this case could increase the cost of green technology adoption and create a profitable market for PAE.

The third chapter explores the technological diversity of Sustainable Development Goals (SDGs) related innovation, focusing particularly on the role of universities. The motivation for this study lies in the fact that much of the literature emphasizes the relationship between interdisciplinary research and the SDGs, which are considered to be closely related; therefore, innovation related to the SDGs is expected to be diverse, especially when it emanates from academia, where teams of researchers collaborate to create innovation for the benefit of society. However, research about innovation for the SDGs is still in its infancy due to the lack of a comprehensive quantitative analysis about its characteristics and the omission of potentially relevant actors, such as universities. This chapter aims to give a threefold contribution analyzing USPTO patent data from 2006 to 2020. First, we develop a novel methodology for tagging SDGs-related patents using an unsupervised natural language processing (NLP) approach; starting from an initial list of keywords, we create an augmented keywords' dictionary for each SDG based on patent text. To do that, we combine the TF-IDF (Term Frequency-Inverse Document Frequency) method with a vectorial representation of patent text and keywords. Second, we analyze universities' contribution to innovation for each SDG. Third, we compare the diversity of SDGs and non-SDGs patents using the Rao-Stirling index to shed light on the interplay between universities, interdisciplinarity and the SDGs. The empirical results point in two directions. On the one hand, SDGs related patents are more diverse than their counterparts across almost all technological fields. However, despite universities' production of innovation related to the SDGs is on the rise, demonstrating an increasing attention to the topic, when we consider university patents only, for most of the SDGs there is not a diversity premium.

Acknowledgements

My doctoral thesis was an intense and fascinating journey. In my opinion, it was a fundamental part of my life that changed me professionally and at the same time made me grow as a person. At almost 29 years old, I can say that I have dedicated at least half of my life to education, and I feel privileged for that. The PhD was the culmination of this long journey, but it was in the latter part of this journey that I learned the most valuable lessons. The first is that your personal motivation plays a big role in enduring the exhausting and nerve-wracking daily life of a PhD student. My motivation is curiosity about the scientific pursuit, which often transcends different fields and disciplines. For this reason, I am deeply indebted to my former and current colleagues and all the professors of the PhD program "Innovation for the Circular Economy" who gave me the opportunity to learn about topics and expertise from different departments and universities. I strongly believe that the current and future challenges we are being asked to address are inherently complex and interconnected, requiring a blend of skills and abilities that result from interdisciplinary education such as that offered by our PhD program.

The second lesson is the importance of independence in becoming a valuable scientist. I would like to thank my PhD advisor Francesco Quatraro for allowing me to pursue my own interests and express my creativity during the PhD, but still guiding me to achieve quality in my research work. In addition, I would like to thank my co-authors Valerio Sterzi, Laura Ciucci, and Claudio Detotto for their patience, kindness, careful review, and scientific guidance, which deeply inspired my research work.

The third lesson is the importance of sharing one's passion and expertise with students. For this, I would like to sincerely thank Professor Laura Corazza and Professor Daniel Torchia, who gave me the opportunity to teach with them in the undergraduate Business Administration program.

Finally, my accomplishments were only possible because of the support of my family. I hope that through my efforts they can be as proud of me as I am of being their daughter.

Contents

List of Figures

List of Tables

Chapter 1

Do universities look like patent trolls? An Empirical Study of University Patent Infringement Litigation in the United States

Joint with Laura Ciucci, Claudio Detotto and Valerio Sterzi

1.1 Introduction

Since the passage of the Bayh-Dole Act in 1980, the number of university patents in the United States has significantly increased (Council et al., 2011) and universities have become increasingly involved in technology transfer activities. At the same time, some universities have also recently become more aggressive in trying to monetize their intellectual property (IP) assets: some studies have found that universities transfer patents to patent trolls (Feldman and Ewing, 2012; Fusco et al., 2019), while others reveal that they are repeat initiators of patent infringement litigation (Rooksby, 2011).

Because universities play an important role in both the production and the dissemination of knowledge, the way they monetize their patents might have a significant impact on society. There are, in particular, specific concerns about the access to university-created inventions that in many cases are publicly funded (Drivas et al., 2017; Thompson et al., 2018).

Despite the importance of this phenomenon, only few studies have examined university patent litigation. Our research partially fills this gap in the literature by providing an empirical analysis of university patent litigation strategies, by comparing patents litigated by universities to those litigated by patent trolls. The reason for this parallel is that both universities and patent trolls are non-practicing entities (NPEs), i.e. entities that do not intend to practice their patents (Lemley, 2007).

Patent trolls are often criticized for harming innovation by enforcing weak and outdated patents that have been infringed by third parties (Chien, 2010). In particular, patent trolls are accused of primarily focusing on suing for patents that relate to an already developed product, thus harming companies and individuals who do not have sufficient resources to fight them in court (Chien, 2013; Feldman and Lemley, 2015; Cohen et al., 2019). Scholars have taken a keen interest in their patent acquisition activities, but especially in their litigation activities, because of the latter's damaging potential on innovation (Bessen and Meurer, 2013*a*; Chien, 2013; Kiebzak et al., 2016). Their influence in the patent market is of great importance as the study by Allison et al. (2010) suggests that patent trolls are the most litigious actor in the market, as almost 50% of the most litigated patents belong to them.

In this regard, universities, as all NPEs, enjoy a privileged position when asserting their patents, such as advantages in litigation tactics, the impossibility of counter-attack by the defendant, and a pure financial interest in the patent (Lemley and Feldman, 2020).

Since 2000, universities are increasingly pursuing patent enforcement (Ascione et al., 2021), sometimes with remarkable success. For example, the California Institute of Technology recently won a $1.1 billion patent verdict against Apple and Broadcom for infringement of patents held by the university on Wi-Fi technology, the largest jury verdict in 2020 and the sixth largest patent verdict ever. Also noteworthy is the University of New Mexico (UNM) case, which in 2019 asserted two patents that were not invented by UNM faculty, but were purchased from the Industrial Technology Research Institute (ITRI), a Taiwanese entity established and funded by the Taiwanese government.

In this paper, we therefore focus on the increasing presence of universities in the courts in the United States. With a few exceptions (Shane and Somaya, 2007; Rooksby, 2011, 2012, 2013; Firpo and Mireles, 2018, 2020), economists and legal scholars have not extensively studied the phenomenon of university patent litigation and the few empirical studies that do address university litigation left unanswered numerous questions about litigation strategies that universities regularly adopt. We thus seek to understand whether the increasing presence of universities in court is related to an increased willingness of universities to defend their intellectual property in order to better ensure its transfer or, rather, to an increased interest in purely monetizing their patents.

We contribute to the debate in several ways. First, we expand the coverage of the phenomenon analyzed by Rooksby (2013); Firpo and Mireles (2018, 2020), both in terms of time span and data considered, i.e. we consider all USPTO patents between 2003 and 2016. Second, we directly compare the characteristics of patents litigated by universities to those litigated by both patent trolls and other entities (which in most cases are product companies). Finally, we deepen our understanding of the phenomenon by analyzing whether the results differ depending on whether the litigated patents belong to the ICT sector, the industry in which trolls litigate the most and where patents are often used for strategic and monetization reasons (O'Neill, 2006; Pohlmann and Opitz, 2013).

The rest of the article is organized as follows. In the next section we present the background literature and the goals of our research. In Section 3, we describe patent and litigation data and we present our variables of interest. In Section 4, we describe the methodology and report the results of our empirical analysis. Section 5 concludes with the discussion of the policy implications of our results.

1.2 Background and motivation

University research is a building block for research and development (R&D) in the most advanced countries. In the US, higher education institutions (HEIs) accounted for around 13% of total R&D expenditure and 50% of basic research in 2017, with the federal government as the largest funder providing more than half of the total R&D budget (Beethika Khan and Okrent, 2020).

Traditionally, universities have been perceived as a support structure for innovation, whose approach to knowledge creation and diffusion influences the entire economy (Nelson, 1993). However, their role in the innovation arena changed after the second academic revolution (Etzkowitz, 2001), when universities started to engage in entrepreneurial activities, not without tensions, in order to transform the outcomes of their inventions into patented and thus marketable products (Etzkowitz, 1990). Therefore, HEIs changed their role in the innovation ecosystem, by integrating university culture with the commercial assumptions underlying IP law (Ghosh, 2016). Therefore, such changes in the relationship between HEIs and intellectual property are interesting to study for the innovation system as a whole (Adams, 1990).

In particular, while there is no doubt about the increasing involvement of universities in patenting activity (Ryan Jr and Frye, 2017), there is no consensus on whether this has effectively promoted technology transfer (Feldman and Lemley, 2015). On the one hand, the increasing involvement of universities in patenting activity has created financial incentives for companies to develop and commercialize products that would otherwise entail untenable trial costs, and for university professors to support the development and commercialization of inventions that are often at an embryonic stage and require further development efforts from their inventors (Jensen and Thursby, 2001; Ouellette and Weires, 2019). On the other hand, IP protection of university inventions has also had some unintended consequences for the diffusion of scientific knowledge. Not only does academic patenting hinder the use and production of science (Murray and Stern, 2007), but patent licensing agreements are sometimes "divorced" from innovation (Feldman and Lemley, 2015) and, when they are part of technology transfer, they appear to be incidental, since actual knowledge transfer takes place much earlier, through collaborations and informal know-how (Thompson et al., 2018). Further, it should be taken into account that knowledge absorption is also related to the firm absorptive capacity (Cohen and Levinthal, 1989). Other types of university technology transfer tend to be generally more important than university-licensed patents, and alliances between universities and private firms can often take place ex ante, even before patents are filed (Martínez and Sterzi, 2019). Notwithstanding the fundamental role of licensing, commercialization of academic patents can also occur through other mechanisms such as science parks and new business formations, including spinoffs and startups. These provide academic entrepreneurs with an alternative route to disseminate and commercialize research, especially when they are unable to license their technology to large corporations or an outside entrepreneur because the technology is still in its early stages (Lowe, 2002). Occasionally, a spin-off or startup is the only viable option for developing a technology. Without the creation of a new company, the technology may never be commercially viable (Shane, 2004). In addition, spin-offs and companies may have the prospect of alternative funding methods to advance their research agenda (Bercovitz and Feldman, 2006). Supportive facilities such as incubators and science/research parks within or near the college facilitate the creation of spinoffs (Heinzl et al., 2013). For instance, Rothaermel and Thursby (2005) emphasize the relevance of university ties for incubator firms. However, their findings suggest that the absorptive capacity of incubator firms is a crucial aspect for the translation of academic knowledge into competitive advantage at the firm level.

In line with the increasing involvement of universities in patenting activities, universities started to establish technology transfer offices (TTOs) to assist researchers with patent applications and manage licensing revenues, especially after the Bayh-Dole Act. However, licensing does not seem to be too fruitful for universities. In the US, revenues from patent licensing activities account for only a fraction of total university research expenditures (Collinsworth and Crager, 2014; Eisenberg and Cook-Deegan, 2018), and few universities have made profits from licensing activities. An extensive study by Brooking Institution (Valdivia, 2013) shows that 84% of US universities did not break even in technology transfer in 2012, given the staffing and filing costs. Moreover, empirical evidence also shows strong heterogeneity across universities: looking at data from 155 universities, the eight universities with highest licensing revenues accounted for 50% of licensing revenue across the entire sample, suggesting that few patents granted to academic institutions are likely to be highly valuable (Ryan Jr and Frye, 2017).

Thus, in the attempt to increase revenue streams from patenting activity, TTOs started to make use of heterogeneous monetization strategies. First, they increasingly use exclusive (rather than non-exclusive) licenses, although these often act as a barrier to downstream R&D rather than an incentive (Mazzoleni, 2006; Özel and Pénin, 2016).[1] Second, universities also increasingly rely on patent markets to sell their patents (Fusco et al., 2019; Caviggioli et al., 2020). For example, USPTO data on patent reassignment in recent years show that universities are among the entities with the highest volume of outbound patent

[1] In this sense, Ayres and Ouellette (2016) suggest that universities should be allowed to propose exclusive licenses only after first "offer[ing] the invention under a nonexclusive license for a nominal fee".

assignment transactions.[2] In addition, there is recent evidence that universities have adopted new ways of patent marketing by organizing auctions (e.g., in the case of Pennsylvania State University) or relying on auctions organized by third parties (for example, in the case of Ocean Tomo[3]), and by collaborating with patent trolls (Feldman and Ewing, 2012; Fusco et al., 2019; Love et al., 2020). For example, Intellectual Ventures, one of the largest patent holding companies in the US and a notorious patent troll, has disclosed its relationships with more than 400 universities (60 of which are American-including Columbia, University of California, Texas University, and California Institute of Technology), although only two of these have resulted in commercial products (Cordova and Feldman, 2015)[4]; Love et al. (2020) confirm that in most instances university patent transfers are made to patent trolls and note that these transfers very often concern legal liability without really encouraging commercialization. Third, recent evidence also suggests that universities have started to enforce patents at the end of their patent term and against companies that have already developed successful commercial products without having actually exploited the university patent (just as patent trolls would do) (Firpo and Mireles, 2020). In 2014, for example, Boston University won settlements with 25 companies (included large companies in the tech industry such as Amazon, Apple and Microsoft, among others) it sued for infringing their patented technology (US5686738 - filed in 1995) for producing blue light-emitting diodes (LEDs).[5] Another famous case of late enforcement of patents was Carnagie Mellon University, which in 2016 received $750 million from Marvell Technology as a settlement for a case of infringement of two patents (US6201839, filed in 1998 and US6438180, filed in 1999) whose purpose was to reduce "noise" on hard drives. This case set the record for the largest payment in a patent case related to a computer science invention (Rooksby, 2016).

The strong interest of universities in patent enforcement became apparent when university associations lobbied against anti-troll legislation that would also thwart their ability to engage in litigation (Valdivia, 2015). In addition, another step appears to be about to be taken by a consortium of 15 of the most prominent US research universities, including California Institute of Technology, Berkeley, Columbia, Harward, and Yale, as explained in a June 10, 2021, EFF article.[6] Indeed, a patent exploitation company has been formed by this consortium with the goal of "receiving payments for patents that have not been successfully licensed via a bilateral 'one patent, one license' transaction". This new LLC will specifically focus on sectors characterized by software patents, as well as non-exclusive sub-licensing. The EFF article

[2] In January and February 2019, four universities appeared among the top ten patent assignors by number of transactions (*IAM Magasine*, 2019).

[3] US university patents account for 20% of business for Ocean Tomo, a company known for its patent auctions (Ledford, 2013).

[4] Recent research by Fusco et al. (2019) confirms the relationship between Intellectual Ventures and universities, showing that the former appears to be the buyer of almost two hundred university patents issued at the USPTO, representing about 50% of all patents transferred to patent trolls.

[5] Incidentally, other companies did not settle and the jury found that they infringed the '738 patent and failed to prove the patent's invalidity. However, the defendants appealed to the Federal Circuit, which reversed and made the patent invalid for not meeting the enablement requirement.

[6] https://www.eff.org/deeplinks/2021/06/15-universities-have-formed-company-looks-lot-patent-troll (accessed July 2021).

concludes that the consortium "will use the threat of litigation to try to get all competitors in a given industry to pay for the same patent".

Concerns about the behavior of some universities in pursuing their litigation strategies have also been shared by US Courts and judges, such that universities are now increasingly considered as for-profit businesses and they are increasingly unlikely to be granted immunity under the experimental use exemption (Rowe, 2005).[7]

Given this background, some scholars have analyzed the consequences of patent litigation in universities (Shane and Somaya, 2007) and questioned to which extent universities can be compared to patent trolls (Lemley, 2007; Firpo and Mireles, 2018, 2020; Rooksby, 2011, 2013). As early as in 2007, Shane and Somaya warned of the potential negative effects of an excessive presence of universities in court. In particular, they focus on the impact that patent litigation can have on universities' efforts to license their technologies. Using an unbalanced panel of 116 universities litigating their patents in court between 1991 and 2000, they find that litigation has a negative impact on licensing activity and suggest that this result is due to litigation disrupting overall TTO's activity by reducing the resources available to commercialize technologies and build licenses.

More recently, Rooksby (2011) and Firpo and Mireles (2018, 2020) describe the characteristics of cases of infringement lawsuits initiated by universities (and the patents at issue), raising questions about the potentially strategic nature of universities' behavior in these litigations. In particular, Rooksby (2011) study includes 57 cases between 2009 and 2010 involving 125 patents. The author concludes that certain characteristics, such as the preference for a jury over a judge, suggest that university behavior is similar to that of for-profit actors. Firpo and Mireles (2018) examine litigation cases filed by universities, foundations, and nonprofits organizations from 2000 to 2015, totaling 585 cases, and find partial evidence of strategic (troll-like) behavior in university patent litigation. However, their analysis is based only on indirect evidence, since they compare the mean values of patent quality for their sample of patents litigated by universities with the mean values of (i) the "most litigated patents" from Allison et al. (2009) and (ii) non-university patents litigated at the Federal District Court of Texas Eastern - both of which are considered evidence of aggressive patent assertion. In a more recent paper, Firpo and Mireles (2020) continue to investigate the troll-like behavior of universities and non-profits, finding inconclusive evidence on the matter but emphasizing that "some universities are engaging in behavior that can be particularly troubling."

In this paper, we build on Firpo and Mireles (2018, 2020) and compare the characteristics of infringement cases initiated by universities with those initiated by patent trolls and other entities in the United States.

[7]In particular, a recent opinion by the Federal Circuit (Madey v. Duke) makes clear that universities should not be granted immunity under experimental use (Rowe, 2005), pointing out that "Duke [...] like other major research institutions of higher learning, is not shy in pursuing an aggressive patent licensing program from which it derives a not insubstantial revenue stream".

In doing so, we analyze three dimensions that have been identified in the literature as characterizing patent trolls' behavior: (i) the intensity with which a patent is litigated (enforcing the patent with a large number of defendants in a short period of time), (ii) the choice of Federal District Court of Texas Eastern to prosecute the litigation, and (iii) the quality of the invention protected by the patent.
Furthermore, we compare these dimensions across two sample of technologically heterogeneous patents: ICT vs non-ICT. This, because scholars agree that patent trolls are primarily active in the ICT industry (Orsatti and Sterzi, 2018) and have an interest in all complex technologies (Kingston, 2001), i.e., those in which a new product or process consists of several separately patentable elements, leading to fragmentation of the respective intellectual property ownership. ICT systems are developed by gradual advancements, with many new technologies building on current technology for reasons of interoperability and cost-effectiveness, increasing the risk of "hold-up" among innovators (Orsenigo et al., 2010). Nonetheless, ICT patents underlying profitable systems are seen as especially luring acquisitions by trolls, who are attracted both by the fragmentation of intellectual property and by the potential multitude of further developments that a single ICT invention could bring (O'Neill, 2006).

Number of defendants

In order to exploit economies of scale, patent trolls often litigate their patents against multiple defendants simultaneously, a habit not shared by producing companies (Allison et al., 2010; Lemley and Feldman, 2020). This strategy is economically feasible because the cost of proving further infringement does not increase linearly since the legal apparatus built up for one defendant can be used against many (Allison et al., 2009, 2010; Feng and Jaravel, 2020). As a result, patent trolls are not afraid to name multiple defendants when they engage in litigation and often seek a quick settlement (Chien, 2013).

Federal District Court of Texas Eastern

Filing suit in Federal District Court of Texas Eastern is widely recognized in the literature as a strategic decision characteristic of troll-like behavior (Rooksby, 2011; Cohen et al., 2016; Firpo and Mireles, 2018; Cohen et al., 2019). Traditionally, the Eastern District of Texas has established itself as a jurisdiction that has attracted a large number of patent infringement cases filed by patent trolls (Cohen et al., 2019) because of its reputation as a favorable venue for plaintiffs in patent infringement actions suits (Masters and Weber, 2009) and because of the expeditious docket (Coursey, 2009). The choice to file suit at this court, when no other ties with the venue are acknowledged, might depend on the idea that the presiding judges sympathize with IP owners (Liang, 2010).

Patent quality

There is quite a heated debate about the quality of patents litigated by patent trolls. On the one hand, empirical evidence shows that patent trolls have lower rates of success compared to producing companies (Allison et al., 2010; Lu, 2012; Risch, 2015; Allison et al., 2017). In particular, data from Darts-IP (2018) show that in Germany, the European country most affected by patent troll-related litigation, producing companies win infringement cases almost 15% more often than patent trolls do. Cohen et al. (2019) find that patent trolls disproportionately assert patents close to their expiration date, which they consider an indication of low-quality lawsuits. On the other hand, Risch (2012) analyzes the patents asserted by the ten most-litigious patent trolls in the United States and found them to be of similar or of higher quality than those asserted by producing companies, when quality is approximated by the number of forward citations or by the number of claims in the patent application.

Thus, in recent years, the literature has already shown how the behavior of patent trolls and producing companies differs with respect to these three dimensions (Allison et al., 2009; Fischer and Henkel, 2012; Firpo and Mireles, 2018; Cohen et al., 2019). In this context, the aim of this paper is to analyze the characteristics of patents litigated by universities in order to understand to what extent they are similar to those litigated by patent trolls.

1.3 Data sources and main variables

1.3.1 Sample construction

We conduct our analysis by relying on US Patent Litigation Docket Reports dataset (PLDR, 2019 version), which contains complete patent litigation information on district court cases filed in US district courts from January 1, 2003, to December 31, 2016, for a total of over 55'000 cases (Schwartz et al., 2019). PLDR provides the information on the type of litigation (infringement actions, patent invalidity, etc.), the names of the parties involved in the litigation, the litigation venue, and the relative litigated patents.

In our analysis we consider only infringement cases where the plaintiff is the patent holder and sues the defendant(s) for infringement of a utility patent, which count 43'663 cases and 33'676 unique patents.

We then enrich the database by identifying infringement cases that involve universities or patent trolls among the plaintiffs. We categorize a litigation case as "**university litigation**" if at least a patent asserted in the infringement case is owned, at the time of the litigation, by a university (TTO included), a research institution, or a hospital. We identify the type of patent owners by relying on EPO PATSTAT Person Augmented Table (EEE-PAT) (Van Looy et al., 2006) that reports for each patent the names of the assignee(s) and its sector (university, public research organization, hospital, company, individual). We then do a match by patent and name of the assignee(s) with US Patent Assignment Dataset (PAD, Version 2017) and, whenever a correspondence was not found, we perform an automatic search for keywords in the assignee name in order to allocate unassigned entities to a specific sector.[8]

Whenever a university was not found, we then verify whether a patent troll is at the origin of the litigation ("**patent troll litigation**"). For this purpose, we follow two strategies. First, we identify patents owned by patent trolls at the time of the litigation by relying on a list of patent troll names (and theirs subsidiaries) provided by Darts-IP.[9] The list contains the names of firms that Darts-IP defines as "independent organizations (legal entities) which own or benefit from patent rights but do not sell or manufacture goods or services associated with them (i.e., non-operating companies) and which have an active (offensive) assertion or litigation role as plaintiffs towards the enforcement of their patent rights".[10] We do a probabilistic match between the assignee names in PAD and the list of patent troll names in order to identify if and when a patent was held by a patent troll. We identify 12'969 litigated patents where a patent troll is among the plaintiff (this corresponds to the category "Identified patent trolls" in Table 1.1). Second, since patent trolls make often use of dormant and shell companies with the purpose of litigating their patents without disclosing patent ownership and reducing personal liability (Morton and Shapiro, 2013; Federal Trade Commission, 2016; Sterzi et al., 2021*a*), we extend the list of identified patent trolls provided by Darts-IP by including all entities that take the form of limited liability company (LLC), which is the most common type of entity used by patent trolls (Sterzi et al., 2021*a*). This is, for example, the case of Intellectual Ventures: Darts-IP identifies almost two hundred entities linked to Intellectual Ventures, while for other sources the overall number of shell companies exceeds two thousands, and practically all of them take the form of LLCs.[11] By including LLC entities, we identify 8'240 additional patents asserted by patent trolls (this corresponds to the category "Other LLC entities" in Table 1.1). For robustness, we will also exclude these patents in some regressions in the empirical analysis.

Finally, all the remaining cases are categorized as "**Other entities**", which consist of infringement suits filed by product companies (in most of the cases) and by individuals. The final database accounts for 43'663 infringement actions filed in the US and 33'676 litigated patents between 2003 and 2016. Table 1.1 shows the frequency for each of the three categories. Infringement cases filed by universities ("University litigation") are still quite a rare occurrence (1.3% of all the cases), while patent trolls file a more significantly number of patent infringement lawsuits as they account for almost 30% or 50% of the cases, depending whether we consider (or not) LLC entities as patent trolls.

[8] For example, we use the business entities code to individuate private business enterprises and keywords like "school" or "university" to identify universities.

[9] https://clarivate.com/darts-ip/ (accessed June 2021)

[10] https://www.darts-ip.com/de/npe-litigation-in-the-european-union-facts-and-figures-2/ (accessed July 2021)

[11] https://www.plainsite.org/tags/intellectual-ventures-shell-companies/ (accessed June 2021)

Table 1.1: Number and share of infringement cases by type of plaintiffs

	Frequency	Percent	Cum
University	579	1.33%	1.33%
Patent troll	21'209	48.57%	49.90%
*Identified patent trolls**	12'969	29.70%	
Other LLC entities	8'240	18.87%	
Other entities	21'209	48.57%	100%
Total	43'663	100.00%	

Notes: Litigation years: 2003-2016. Infringement actions only.
* Identified patent trolls refers to the Darts-IP list of trolls

For each litigation case, we then collect the information on the litigated patent, such as the technological class and the application year, from the US-OECD patent quality database (version 2020) (Squicciarini et al., 2013) and the forward citations from the Patents View Database[12].

1.3.2 Main variables

In line with the characteristics highlighted at the end of Section 1.2, litigated patents are analyzed on three main dimensions to identify similar patterns across groups. The first characteristic under study is the number of defendants per patent ($DEFENDANTS$). The number of defendants refers to the sum of all the defendants by litigation case. Second, a binary variable has been created taking the value of one when a litigation is filed at the Federal District Court of Texas Eastern ($TEXAS$). A relevant share (about 1 out of 4 cases) of infringement actions takes place at this Court that, as explained before, has earned the reputation of being friendly to patent holders and has attracted numerous opportunistic patent litigations (Cohen et al., 2016). The third dimension accounts for the quality of the patent that we proxy in two ways. First, following Cohen et al. (2019), we consider patents litigated late in time as a sign of low litigation value. In particular, we consider the age of the patent at the time of the litigation (AGE), where age is defined as the time lag between grant date and litigation date. Second, we consider the number of citations received by the focal patent as they are indication that an innovation has contributed to the development of subsequent inventions (Henderson et al., 1998). In doing so, we build two measures that differentiate the technological importance of the patent at the time of the invention from the quality at the time of the litigation. The first measure ($5Y\_FILING\_CITATIONS$) indicates the number of citations received by the focal patent in the first five years after the filing date; the second measure ($5Y\_LITIGATION\_CITATIONS$) indicates the number of citations computed in the year of the litigation and in the four years before. While the former is largely used in the literature, measuring the potential technological and commercial importance of a patent at the beginning of its lifetime, the

[12] https://patentsview.org/ (accessed June 2021)

latter has the benefit of representing the economic value of a patent at the time of the litigation.

Due to missing data on some covariates, the dataset reduces to 87,919 records, which corresponds to 41,191 infringements and 30'882 litigated patents over the time span 2003-2016 in the US. Table 1.2 depicts the descriptive statistics of the dataset under study. This descriptive analysis provides evidence of differences and/or similarities across the three groups, namely Patent trolls, Universities and Other entities, considering the main characteristics of litigated patents, i.e. number of defendants, venue of litigation (Texas) and patent quality (patent age, the number of citations at the time of the invention, and at the time of the litigation).

Overall, taking the full sample, sharp differences are observed between patent trolls and other entities, where the former exhibit higher values in all respects. Dissimilarities look a little milder between trolls and universities, where the litigated patents are similar for age at the time of the litigation. Moreover, when looking at ICT sector, differences in litigated patents are smaller among patent trolls and universities. First, universities litigate a significant share of their patents in Texas (22%), below 40% of patents litigated by patent trolls, but above 16% of patents litigated by other entities. In addition, no more difference is found with respect to the number of defendants per patent and the number of citations at the time of the litigation.

The next section provides a more analytical look at the differences among the three identified groups of patent plaintiffs and it aims to propose a more detailed model that builds on these intuitions. To do so, an econometric model is estimated that enables to isolate the relationship between patent characteristics and group belonging. We employ a multiple regression framework since it retains the independent variables in their original form, thus making model interpretation easier than alternative multivariate procedures, such as Linear discriminant analysis or Multivariate analysis of variance.

1.4 Empirical Analysis

As explained before, this section is devoted to study the litigation strategies of plaintiffs by providing a broad-based statistical characterization of patent cases filed by different typologies of plaintiff, namely patent trolls, universities, and other entities. In particular, looking at the main characteristics of the patents under litigation over time, the aim of this empirical study is to identify some regularities, if any, in the recent increase in university patent litigation and to explore evidence of strategic behavior in order to identify similarities and differences across types of plaintiffs. The analysis concludes with a focus on *ICT*, i.e. the field that is targeted the most by patent trolls (Fischer and Henkel, 2012).

Table 1.2: Descriptive statistics (2003-2016) by assignee typology and sector

Variables	mean	sd	OTHER mean (A)	TROLLS mean (B)	UNIVERSITIES mean (C)	(1) t-test	(2) t-test	(3) t-test
				Full sample				
DEFENDANTS	5.12	11.71	4.79	5.48	4.40	-8.64***	1.39	-3.17***
TEXAS	0.20	0.40	0.09	0.32	0.07	-87.75***	2.09**	20.07***
AGE	10.02	5.34	9.26	11.27	11.43	-56.43***	-15.99***	-1.07
5Y_FILING_CITATIONS	14.87	24.90	12.78	17.10	11.62	-25.47***	1.96**	7.62***
5Y_LITIGATION_CITATIONS	20.97	37.68	17.95	23.99	22.06	-23.59***	-4.54***	1.76*
N. Obs.	87'919		43'580	42'889	1'450			
				ICT				
DEFENDANTS	5.96	13.92	5.80	6.03	6.56	-1.71*	-1.01	-0.67
TEXAS	0.31	0.46	0.16	0.40	0.22	-55.15***	-3.21***	6.47***
AGE	10.96	5.45	9.61	11.69	11.43	-40.05***	-6.53***	0.84
5Y_FILING_CITATIONS	18.99	30.28	19.21	18.92	14.21	0.99	2.94**	2.84***
5Y_LITIGATION_CITATIONS	25.80	42.70	25.37	26.06	22.87	-1.69*	1.06	1.34
N. Obs.	47'334		16'520	30'481	333			
				non-ICT				
DEFENDANTS	4.15	8.33	4.18	4.13	3.76	0.50	1.67*	1.36
TEXAS	0.07	0.26	0.05	0.13	0.03	-29.07***	2.85***	9.89***
AGE	9.47	5.09	9.04	10.24	11.42	-21.95***	-15.31***	-7.74***
5Y_FILING_CITATIONS	10.07	15.20	8.86	12.63	10.84	-22.96***	-4.70**	3.29***
5Y_LITIGATION_CITATIONS	15.33	29.84	13.43	18.90	21.82	17.00***	10.30***	-2.64***
N. Obs.	40'585		27'060	12'408	1'117			

Notes: (1), (2) and (3) represent, respectively, the t-values of the tests on the difference between the following sample means: (A)-(B), (A)-(C) and (B)-(C). ***p<0.01,**p<0.05, *p<0.1.

1.4.1 Patent trolls

In our first econometric exercise, we first characterize the litigation strategies of patent trolls by estimating the likelihood that a given litigated patent is asserted by a patent troll rather than another entity. To do so, we estimate the following Logit model in which the dependent variable is a dichotomous variable whereby litigated patents belonging to patent trolls are designated as "1" and "0" otherwise:

$$Pr(Y_i = 1) = \frac{e^{\beta' x}}{1 + e^{\beta' X}} \tag{1.1}$$

In other words, the probability to observe a patent litigated by a patent trolls - rather than other entities - is assumed to be a function of a vector of explanatory variables, x, namely the three dimensions as described in the previous sub-section.

The results are shown in Table 1.3 and confirm the hypothesis in Section 1.2: patent trolls litigate their patents more against a higher number of defendants and more frequently in Texas. In particular, a patent litigated in Texas has 137% more odds of being litigated by a troll compared to patent not litigated in Texas. Further, an additional defendant is associated with 0.5% higher odds, while an additional year of age with

Table 1.3: Patent troll characteristics. Logit regression results

VARIABLES	(1) TROLL (=1) vs NO-TROLL (=0)
DEFENDANTS	0.00532***
	(3.235)
TEXAS	0.864***
	(18.77)
AGE	0.0322***
	(11.77)
5Y_FILING_CITATIONS	0.00180***
	(3.039)
5Y_LITIGATION_CITATIONS	-0.00130***
	(3.255)
Constant	-3.058***
	(19.12)
Observations	87'919
Year dummies	YES
Field dummies	YES
Pseudo R2	0.247

Notes: Unit of observation: patent. Litigation years: 2003-2016; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

3.27% higher odds of being litigated by a troll. With respect to the quality, patent trolls initiate litigation on relatively high quality patents when referring to a standard measure ($5Y\_FILING\_CITATIONS$): patents litigated by patent trolls receive more citations at the beginning of their lifetime than the average. However, the probability of a patent being litigated by a troll also decreases with the technological importance at the time of litigation ($5Y\_LITIGATION\_CITATIONS$): an additional citation in the year of litigation or in the four before, decreases the odds of being litigated by a troll by 0.13%. In other words, these results, together with the positive coefficient of AGE, suggest that trolls tend to litigate patents that are of high quality at the time of filing but that are less important at the time of litigation. Altogether these results corroborate the findings that patent trolls mainly monetize valuable but old technologies rather than playing as intermediaries in the patent market (Orsatti and Sterzi, 2018).

1.4.2 Comparing universities to patent trolls

In what follows, we want to analyze the patent litigation strategies of universities by comparing the characteristics of the patents they litigate with those that are instead litigated by patent trolls. In doing so, we estimate Multinomial Logit models that explore the factors explaining the likelihood of a patent belonging to one of the three groups under study, namely Trolls, Universities and Other entities:

$$Pr(y = j) = \frac{e^{\beta_j x}}{1 + \sum_{k=1}^{J-1} e^{\beta_k x}} \qquad for \quad j = 1, 2, ..., J - 1 \tag{1.2}$$

Where $Pr(y = j)$ is the probability of the n-th patent belonging to the j-th group and x is a vector of explanatory variables. Other entities group (OTHER) is designated as the reference category, J. The probability of membership in other categories is then compared to the probability of membership in the reference category.

The results of the Multinomial Logit regression are shown in Table 1.4. Columns (1) and (2) provide the outputs corresponding to TROLL and UNIVERSITY equations, while column (3) gives the statistical tests on the differences between coefficients across the two equations. First, it can be noted that all the coefficients in column (1) are significant and consistent with those obtained with the Logit model (Table 1.3). Second, contrary to patent trolls, the characteristics of patents litigated by universities are more similar to those litigated by the reference group: university litigated patents differ significantly from the reference group only on two variables (AGE and $5Y\_LITIGATION\_CITATIONS$). With respect to age, we observe that, as for patent trolls, universities tend to initiate litigation on relatively older patents than those of the reference group. If patent age increases by one unit, the relative log odds of being in trolls and universities group increase by 0.03 and 0.05, respectively. The difference between the two coefficients is small but statistically significant at the five percent significance level. The variable $TROLLS$ has a positive and significant effect for patent trolls: the log odds of being in the troll group for patents litigated in Texas are predicted to be 0.87 points greater than that for patents not litigated in Texas. However, there is no difference between university and the other group. The difference between the coefficient of trolls and universities is significant at one percent significance level. The variable $5Y\_LITIGATION\_CITATIONS$ shows a positive and significant coefficient for universities. In this respect, universities differ from both trolls and other entities by litigating patents of relatively higher quality at the time of litigation. In particular, for an additional citation in the year of litigation and in the four years before, the log odds of a patent being in the university group is predicted to increase by 0.002 units, while that of being in the troll group is predicted to decrease by 0.001 units.

Table 1.5 provides the outcomes of the Multinomial Logit regression on patent ownership group when restricting to ICT sample. For what concerns patent trolls, with the exception of $5Y\_FILING\_CITATIONS$ which is no longer significant, all the coefficients in column (1) are once again significant with the expected sign and a higher magnitude than observed in Table 1.4. This tends to confirm a particularly aggressive behavior of patent trolls in this sector. With respect to $TEXAS$, we observe that, as for patent trolls, universities tend to initiate litigation in Texas more than in the reference group. If patent is litigated in

Table 1.4: Multinomial Logit regression results. Full sample

VARIABLES	(1) TROLL	(2) UNIVERSITY	(3) χ^2 test
DEFENDANTS	0.00531***	-0.000748	0.61
	(3.200)	(0.0960)	
TEXAS	0.872***	0.229	11.76***
	(18.72)	(1.210)	
AGE	0.0340***	0.0553***	4.93**
	(12.37)	(5.826)	
5Y_FILING_CITATIONS	0.00175***	-0.000503	1.62
	(2.906)	(0.281)	
5Y_LITIGATION_CITATIONS	-0.00120***	0.00280**	12.26***
	(2.923)	(2.440)	
Constant	-2.477***	-6.040***	
	(15.41)	(13.79)	
Observations	87'919		
Year dummies	YES		
Field dummies	YES		
Pseudo R2	0.244		

Notes: Unit of observation: patent. Litigation years: 2003-2016; # The group OTHER is the baseline comparison group; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

Texas, the relative log odds of being in trolls and universities group increase by 0.959 and 0.449 units, respectively. The difference between the two coefficients is statistically significant at the five percent significance level. In addition, universities litigate patents of lower quality with respect to the reference group, in particular for an additional citation if the first five years after the filing of the patent, the log odds of being in the university group is predicted to decrease by 0.006 units.

Table 1.6 shows the results of the Multinomial Logit model for the non-ICT sample. The main findings are the same as in the Table 1.4 for the whole sample, with two exceptions. First, for patent trolls and universities, the coefficient related to the number of defendants is not significantly different from zero. The number of defendants per patent is the same across groups. Second, contrary to the ICT sector, patents litigated by universities are of high quality as measured by the number of forward citations; in particular, $5Y\_FILING\_CITATIONS$ now has a positive and significant effect and the difference between the coefficients in the two columns is not statistically significant; further, $5Y\_LITIGATION\_CITATIONS$ has a positive and significant coefficient as well: an additional citation increase the relative log odds of being in the university group by 0.002; on contrary, it decreases the relative log odds of being in the trolls group by 0.002. The difference between the two coefficients is statistically significant at the one percent significance level.

Table 1.5: Multinomial Logit regression results. ICT sample

VARIABLES	(1) TROLL	(2) UNIVERSITY	(3) χ^2 test
DEFENDANTS	0.00669***	0.00562	0.04
	(3.181)	(0.972)	
TEXAS	0.959***	0.449**	5.27**
	(16.96)	(1.997)	
AGE	0.0469***	0.0338**	0.69
	(12.58)	(2.143)	
5Y_FILING_CITATIONS	0.000332	-0.00636*	3.32*
	(0.470)	(1.719)	
5Y_LITIGATION_CITATIONS	-0.00123**	0.00147	2.00
	(2.476)	(0.767)	
Constant	-2.623***	-7.641***	
	(10.76)	(9.735)	
Observations	47'334		
Year dummies	YES		
Field dummies	YES		
Pseudo R2	0.195		

Notes: Unit of observation: patent. Litigation years: 2003-2016; # The group OTHER is the baseline comparison group; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

As a robustness check, the Multinomial Logit analysis is also performed using (i) only the Darts-IP list to identify the patent trolls (Tables 1.8, 1.9 and 1.10) and (ii) only US universities (Tables 1.11, 1.12 and 1.13). The results are presented in the Appendix and confirm the main results obtained in this section.

Table 1.6: Multinomial Logit regression results. Non-ICT sample

VARIABLES	(1) TROLL	(2) UNIVERSITY	(3) χ^2 test
DEFENDANTS	0.00311	-0.0103	0.45
	(1.073)	(0.521)	
TEXAS	0.564***	-0.244	4.53**
	(6.896)	(0.642)	
AGE	0.0177***	0.0602***	10.99***
	(4.422)	(4.826)	
5Y_FILING_CITATIONS	0.0104***	0.00897***	0.23
	(6.770)	(3.044)	
5Y_LITIGATION_CITATIONS	-0.00243***	0.00259*	9.44***
	(2.876)	(1.681)	
Constant	-2.783***	-5.732***	
	(16.68)	(8.349)	
Observations	40'585		
Year dummies	YES		
Field dummies	YES		
Pseudo R2	0.179		

Notes: Unit of observation: patent. Litigation years: 2003-2016; # The group OTHER is the baseline comparison group; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

1.5 Concluding remarks

Over the past twenty years, TTOs have evolved from pursuing patent protection and licensing innovation to elaborating careful commercialization strategies, as revenue generation is now a relevant goal of their technology transfer operations. Indeed, recent evidence shows that university monetization activities have increased significantly, especially in the US, and that university TTOs appear now to be "patent-centric" in their attempt to fit entrepreneurship and commercialization into universities' missions (Carter-Johnson, 2020) as well as being "revenue-driven with a single-minded focus on generating licensing income" (Kesan, 2008).

Currently, some universities are redoubling their efforts to pave the way for sharing and selling their scientific results. For example, in mid-2018, Stanford University reorganized its TTO under a new director, centralizing its functions and hiring new business development staff, to provide “a higher return on marketing efforts”.[13]

While the growing attention to monetization activities might help universities to increase funding, there is no consensus on whether this effectively promotes technology transfer. To be effective, technology transfer should include not only the public information disclosed in the patent, but also the transfer of know-how. However, this is not the case for most of licensing requests from universities, which, on the contrary, are often justified by the need to monetize the patented inventions (Lemley and Feldman, 2020). In this sense, the behavior of universities as non-practicing entities, in dealing with the patent system might be in many ways more akin to patent trolls than to product companies. Moreover, if the purpose of the Bayh-Dole act is to promote commercialization of academic inventions, the increasing participation of universities in litigation activity may be controversial because it is not the university as patentee but the defendant that achieves the goal of the patent system (Lemley and Feldman, 2016).

Our study provides the first comprehensive evidence of the characteristics of litigation strategies of universities in the United States in the time span 2003-2016, by comparing patents litigated by universities to those litigated by patent trolls. Our findings can be summarized as follows. On the one hand, universities seem not to engage in opportunistic litigation as patents litigated by universities differ from those litigated by patent trolls in several respects: unlike patent trolls, universities do not file most of their patent lawsuits in the Eastern District of Texas, do not litigate their patents against multiple defendants, while they litigate highly-cited patents at the time of the litigation. However, we also observe a great deal of heterogeneity at the sector level and a different picture in the ICT industry, the industry where trolls litigate the most and where patents are often used for strategic reasons. In the ICT field, the

[13] https://hechingerreport.org/think-universities-are-making-lots-of-money-from-inventions-think-again/ (accessed July 2021)

characteristics of patents litigated by universities are more similar to those litigated by patent trolls than those litigated by other entities: universities frequently litigate their patents in the Eastern District of Texas and they assert patents that are not of high quality compared to other entities.

Although our results are based on a still-emerging phenomenon, they call for serious consideration of the possible consequences that the rent-seeking behavior may have on technology transfer and innovation, and the evolution of the phenomenon therefore needs to be carefully monitored. Recent initiatives, such as the the Reclaim Invention Program by the Electronic Frontier Foundation (EFF)[14], are to be welcome as they might curb the increasing rent-seeking behavior that characterizes some universities in the United States.

[14] https://www.eff.org/it/reclaim-invention/pledge (accessed June 2021)

1.6 Appendix for Chapter 1

Table 1.7: Correlation matrix among variables under study

	DEFENDANTS	TEXAS	AGE	CITATION
TEXAS	0.049			
AGE	0.011	0.118		
CITATION	0.033	0.105	0.106	
EARLY_CIT	0.046	0.086	0.263	0.721

Table 1.8: Multinomial Logit regression results. Full sample
LLCs excluded from the analysis

VARIABLES	(1) TROLL	(2) UNIVERSITY	(3) χ^2 test
DEFENDANTS	0.00683***	-0.000303	1.01
	(3.446)	(0.0426)	
TEXAS	1.020***	0.227	17.79***
	(18.61)	(1.207)	
AGE	0.0496***	0.0577***	0.67
	(14.60)	(6.055)	
5Y_FILING_CITATIONS	0.00210***	-0.000492	1.65
	(2.968)	(0.241)	
5Y_LITIGATION_CITATIONS	-0.000678	0.00287**	8.69***
	(1.480)	(2.387)	
Constant	-3.178***	-6.087***	
	(12.65)	(13.85)	
Observations	73'474		
Year dummies	YES		
Field dummies	YES		
Pseudo R2	0.360		

Notes: Unit of observation: patent. Litigation years: 2003-2016; # The group OTHER is the baseline comparison group; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

Table 1.9: Multinomial Logit regression results. ICT sample
LLCs excluded from the analysis

VARIABLES	(1) TROLL	(2) UNIVERSITY	(3) χ^2 test
DEFENDANTS	0.00713***	0.00606	0.04
	(3.248)	(1.124)	
TEXAS	1.062***	0.454**	7.71***
	(17.31)	(2.058)	
AGE	0.0606***	0.0354**	2.59
	(14.99)	(2.270)	
5Y_FILING_CITATIONS	-0.000406	-0.00651*	2.84*
	(0.556)	(1.784)	
5Y_LITIGATION_CITATIONS	-0.00104**	0.00140	1.66
	(2.031)	(0.736)	
Constant	-3.488***	-7.735***	
	(9.625)	(9.832)	
Observations	39'538		
Year dummies	YES		
Field dummies	YES		
Pseudo R2	0.224		

Notes: Unit of observation: patent. Litigation years: 2003-2016; # The group OTHER is the baseline comparison group; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

Table 1.10: Multinomial Logit regression results. Non-ICT sample
LLCs excluded from the analysis

VARIABLES	(1) TROLL	(2) UNIVERSITY	(3) χ^2 test
DEFENDANTS	0.00542	-0.00784	0.59
	(1.264)	(0.464)	
TEXAS	0.832***	-0.234	7.54***
	(7.235)	(0.609)	
AGE	0.0248***	0.0622***	7.65***
	(4.020)	(4.982)	
5Y_FILING_CITATIONS	0.0241***	0.00957***	17.68***
	(10.26)	(3.201)	
5Y_LITIGATION_CITATIONS	-0.00405***	0.00281*	11.98***
	(3.047)	(1.697)	
Constant	-3.476***	-5.754***	
	(16.03)	(8.372)	
Observations	33'936		
Year dummies	YES		
Field dummies	YES		
Pseudo R2	0.371		

Notes: Unit of observation: patent. Litigation years: 2003-2016; # The group OTHER is the baseline comparison group; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

Table 1.11: Multinomial Logit regression results Full sample
US Universities

VARIABLES	(1) TROLL	(2) US UNIVERSITY	(3) χ^2 test
DEFENDANTS	0.00531***	0.00337	0.08
	(3.197)	(0.485)	
TEXAS	0.874***	0.174	11.84***
	(18.75)	(0.852)	
AGE	0.0339***	0.0837***	31.14***
	(12.32)	(9.483)	
5Y_FILING_CITATIONS	0.00179***	0.00262*	0.33
	(2.960)	(1.832)	
5Y_LITIGATION_CITATIONS	-0.00124***	0.00201*	8.71***
	(3.027)	(1.838)	
Constant	-2.478***	-6.820***	
	(15.41)	(12.29)	
Observations	87'489		
Year dummies	YES		
Field dummies	YES		
Pseudo R2	0.245		

Notes: Unit of observation: patent. Litigation years: 2003-2016; # The group OTHER is the baseline comparison group; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

Table 1.12: Multinomial Logit regression results. ICT sample
US Universities

VARIABLES	(1) TROLL	(2) US UNIVERSITY	(3)
DEFENDANTS	0.00671***	0.00950*	0.38
	(3.183)	(1.919)	
TEXAS	0.960***	0.255	8.87***
	(16.97)	(1.068)	
AGE	0.0468***	0.0524**	0.07
	(12.57)	(2.520)	
5Y_FILING_CITATIONS	0.000322	-0.00208	0.48
	(0.456)	(0.598)	
5Y_LITIGATION_CITATIONS	-0.00123**	0.000482	0.53
	(2.477)	(0.204)	
Constant	-2.622***	-7.842***	
	(10.75)	(9.375)	
Observations	47'225		
Year dummies	YES		
Field dummies	YES		
Pseudo R2	0.195		

Notes: Unit of observation: patent. Litigation years: 2003-2016; # The group OTHER is the baseline comparison group; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

Table 1.13: Multinomial Logit regression results. Non-ICT sample
US Universities

VARIABLES	(1) TROLL	(2) US UNIVERSITY	(3)
DEFENDANTS	0.00311	-0.00617	0.22
	(1.069)	(0.312)	
TEXAS	0.565***	-0.132	2.90*
	(6.907)	(0.323)	
AGE	0.0176***	0.0934***	55.41***
	(4.373)	(9.407)	
5Y_FILING_CITATIONS	0.0106***	0.0135***	1.03
	(6.859)	(4.784)	
5Y_LITIGATION_CITATIONS	-0.00251***	0.00152	6.80***
	(2.985)	(1.069)	
Constant	-2.785***	-6.503***	
	(16.69)	(11.28)	
Observations	40'264		
Year dummies	YES		
Field dummies	YES		
Pseudo R2	0.178		

Notes: Unit of observation: patent. Litigation years: 2003-2016; # The group OTHER is the baseline comparison group; +Robust z-statistics (in absolute value) in parentheses; ***p<0.01,**p<0.05, *p<0.1

Chapter 2

Patent infringement and public procurement: the effect of environmental stringency on the litigation strategies of green technologies

Joint with Francesco Quatraro and Valerio Sterzi

2.1 Introduction

There is nowadays widespread consensus on the urgency to cope with the dramatic effects of climate change. According the United Nations Environment Programme (UNEP), to keep global warming below the 1.5 °C threshold, polluting emissions have to be reduced by about 7.6% per year to 2030. (Christensen and Olhoff, 2019).

These concerns have catalyzed attention of both policymakers and academic scholars. An increasing body of economic literature has accordingly dealt with theoretical and empirical studies on the decoupling of economic growth from environmental degradation (Meadows et al., 1972; Ayres, 2008). Within this context, particular emphasis has been given to innovation and new technologies as a lever to achieve more sustainable production and consumption, improving environmental performances in terms of both resources exploitation and harmful emissions. The term *eco-innovation* has been introduced to label this kind of innovations (Kemp, 2010), which according to the Porter's hypothesis might have the joint effect of improving the environmental performances of economic activities and of increasing their technical and economic efficiency (Porter and Van der Linde, 1995).

The generation and adoption of eco-innovation are therefore deemed to be crucial for the greening of the economy. Yet, from an economic viewpoint, a major problem is represented by the so-called *double externality* problem, which makes public intervention necessary to restore optimal levels of investment in this specific technological domain (Rennings, 2000). Environmental regulation has been proposed as a means to trigger the diffusion of eco-innovations, which exploits the traditional inducement dynamics aiming at saving on production costs (Johnstone et al., 2012; Ghisetti and Quatraro, 2013; Crespi et al., 2015).

An important implication of these dynamics concerns the transmission of inducement mechanisms upstream to technology suppliers. In fact, the regulatory push-pull effect of environmental regulation not only triggers the adoption of these technologies, but also stimulate the commitment of private resources to their generation, due to the increase of their derived demand and the creation of new market niches, and to the concurrent allocation of public resources to strengthen the knowledge base in this domain (Rennings, 2000; Nemet, 2009; Hoppmann et al., 2013).

Yet, these dynamics have been substantially underestimated by the literature, which has instead much focused on the economic and environmental gains associated to the adoption of green technologies, while devoting little attention to the strategic response of profit-seeking green technology suppliers (Barbieri et al., 2016). Few exceptions in this respect have proposed firm-level analyses of the impact of the production of green technologies on producers' economic performances, proxied by productivity, sales growth and market value, showing that indeed there is as positive effect which is also channeled by the exposure to environmnetal regulation (Marin, 2014; Leoncini et al., 2019; Gagliardi et al., 2016; Colombelli et al., 2020, 2021).

This paper aims at contributing to this strand of literature, by shedding light on the interplay between the regulation-induced prospects for profits in the markets for green technologies and private agents' strategies to reap them, based on patent ligation dynamics. In particular, in the last decade, the literature on the so-called *patents trolls*, has gained momentum. Patent trolls are a kind non-practicing entity (NPE), a company or a person who own patents but does not create or sell products or services. Most definitions further specify that a patent troll must be an entity that exists to assert patents against other actors, birthing the term "patent assertion entity" (PAE) (Miller, 2018). Hence, PAE adopt a strategic (and opportunistic) management of IPRs, by threatening legal prosecution of patents' infringers, to appropriate profits through ex-ante market agreements or ex-post damage awards (Fischer and Henkel, 2012; Reitzig et al., 2007).

A large stream of literature, starting from the seminal contribution by Lanjouw and Schankerman (2001), has stressed that the probability to litigate patents is correlated to patents' characteristics. Similar results have been obtained by specifically looking at litigations started by NPE (Reitzig et al., 2007; Shrestha, 2010). For example, and quite relevant in the green domain, patents covering high-quality inventions on average are more likely to be litigated (Allison et al., 2003; Fischer and Henkel, 2012; Reitzig et al., 2010). However, Lanjouw and Schankerman (2001) find also that an increase in patent breadth is associated with lower likelihood of litigation. This can be explained by the difficulty of the patent holder to detect infringement in case the invention is related to various technological domains (Khachatryan and Muehlmann, 2019).

Accordingly, following empirical evidence on green technologies that emphasize their higher value (del Río González, 2009; Barbieri et al., 2020), we put forth the hypothesis that green patents are more likely to be litigated. Our results, based on the the Stanford NPE Litigation Dataset and the Patent Litigation USPTO Dataset, do not provide support to this hypothesis, showing that green patents have, in general, a lower probability to be litigated, thus corroborating the hypothesis of Lanjouw and Schankerman (2001) for patents with a greater technological breadth. However, when they are owned by NPE, green patents have a greater probability of being litigated. Moreover, the presence of green public procurement (GPP) affects the previously mentioned findings in an interesting way: in the case of green patents litigated by PAE, the presence of GPP increases the probability of litigation. These results have important policy implications, as these dynamics not only are likely to hinder innovation by lowering firms' incentives to allocate resources to R&D, but also lead to underprovision of technologies that could favor the improvement of the environmental impact of economic activities.

This paper adds to the literature in several respects. First, analyses of specific domains can be found in the patents trolls literature, yet no systematic investigations of the green domain have been carried out. Second, the large number of studies on environmental innovations and green patents so far has not addressed the implications of patents trolls strategies. As anticipated, this can yield important indications for policymakers and stakeholders. Third, we elaborate on a possible dark side of environmental regulation when patents trolls are at stake.

The rest of the paper is organize as it follows. Section 2 discusses the background literature and articulates the research question. Section 3 illustrates the research design and the empirical strategy. Section 4 presents the results of the econometric estimations. Section 5 provides the discussion, the conclusions and the relevant policy and managerial implications.

2.2 Literature and Hypotheses Development

2.2.1 Green patenting, policy and economic incentives

Innovation dynamics are deemed to be crucial to cope with the ecological transition. Accordingly, a wide body of literature has investigated drivers and incentives to the introduction of eco-innovation. As stressed by Rennings (2000), the sole economic incentives might lead to sub-optimal investments, because of the so-called *double externality* problem. In this context, on the one hand eco-innovation is affected by the classical source of externality stemming from the quasi-public nature of knowledge (Arrow, 1962). On the other hand, it generates one more collective benefit related to the improvement of environmental conditions, which is not paid for by the market. Consequently, governmental action is seen as indispensable for achieving optimal investment levels. Therefore, the core tenet of Porter's environment and competitiveness hypothesis is the favorable impact of environmental regulation on innovation (Porter and van der Linde 1995). The debate in this context revolves around the need for a mix of measures ensuring the appropriate economic incentives to the generation and adoption of green innovations (Crespi and Quatraro, 2013, 2015; Crespi et al., 2015; Costantini et al., 2015). On the basis of the *double externality* problem, governments are expected to play a significant role in acquiring the necessary knowledge in this area; therefore, they must consider how to both support technology development (supply push) and create markets (demand pull) for environmental technologies (Norberg-Bohm, 2000). On the one hand, supply-side policy instruments promote the development of technological capabilities in green sectors via specialized R&D programs (Costantini et al., 2015). On the other hand, demand-side policy instruments entail setting pollution limits to encourage firms to upgrade their production processes, so as to enhance the environmental performance or setting technological standards that would affect prices (Orsatti, Perruchas, Consoli and Quatraro, 2020). This inducement effect is the result of a comparison between the costs that firms would suffer if they continue to pollute and the costs associated with technical upgrading under the new regulations (Ghisetti and Quatraro, 2013). Consequently, a stricter regulatory framework creates an inducement effect that leads firms to have more incentives to adopt organizational and technical innovations to comply with regulations. Eventually, the increased demand for green technologies leads to an expansion of existing markets, which creates further upstream incentives to invest in green R&D (Testa et al., 2012). This study focuses on green public procurement, a demand-side policy tool that has not received much attention in the context of green technologies, although it can leverage green markets through the significant dimension of public purchases, thus accelerating the transition to sustainable growth (Cheng et al., 2018). GPP is considered a "market trigger" because of the indirect effects on both product development and consumer demand for green products (Parikka-Alhola, 2008; Testa et al., 2012). In addition, GPP can contribute to environmental management approaches and eco-design solutions by supporting the closure of material loops through reused materials or the elimination of certain harmful substances to meet buyers' interests (Günther and Scheibe, 2006). Ghisetti and Quatraro (2017) and Orsatti, Quatraro and Pezzoni (2020) provide specific studies on the topic. In particular, Ghisetti and Quatraro (2017) emphasizes that innovative public procurement could foster green innovation by activating a demand-pull dynamic that creates new market niches and thus promotes the diffusion of these technologies. Orsatti, Quatraro and Pezzoni (2020) study the interplay between GPP and the composition of capabilities to stimulate local environmental innovation capacity; their results show the positive relationship between GPP and the generation of green technologies, especially of mitigation technologies.

2.2.2 Green patents litigation strategies and hypotheses development

The United States and many other countries protect inventors' intellectual property through patents, property rights that give the owners of inventions exclusive commercialization or exclusion rights (the right to prevent others from using or selling similar ideas) for a limited period of time. Clearly defined property rights are a feature of well-functioning markets, and patent holders combat alleged infringers primarily through legal action (or threat of legal action). On the one hand, the total number of patent lawsuits has increased by a factor of 10 since 2000 (Cohen et al., 2016). While there may be multiple causes for this increase, there is a body of work that links certain features of patents to the likelihood of being litigated. Many relevant researchers, starting with Lanjouw and Schankerman (2001), have linked better patent quality to a higher risk of being litigated. The authors find that a 10 percent increase in the number of claims is associated with a 1.4 percentage point increase in the likelihood of being litigated, and that one additional forward reference per claim increases the likelihood of an infringement suit by 22 %. Other research supports these findings: according to Allison et al. (2003), litigated patents are valuable patents, as they contain more claims, forward citations, and backward citations. In a later work, Allison et al. (2009) identified the most litigated patents and found that they are characterized by a higher market value. Green patents have been proved to be of greater value in terms of novelty and quality (Messeni Petruzzelli et al., 2011; Barbieri et al., 2020). Furthermore, Dechezleprêtre et al. (2013) show that green technologies are much more cited than the others, as they receive citations from both green and non-green patents. Consistently with the above literature , we put forward the following hypothesis:

Hypothesis 1 (H1): *Green patents, being of higher quality, are at a greater risk of being litigated*

On the other hand, the percentage of patent lawsuits brought by patent trolls increased from 35% to 69% between 2010 and 2015 (RPX, 2015; CEA, 2016). It has been argued that rather than promoting innovation, the patent system increasingly hinders it by forcing companies to engage in costly patent litigation (Bessen, 2005; Kiebzak et al., 2016; Smeets, 2014). These costs include monitoring and detection, legal advice, claim creation, negotiations, and potential damages (Bessen and Meurer, 2008). Moreover, the notion that stronger patent rights promote innovation depends on the premise that patent enforcement is successful enough to effectively solve legal issues. In fact, the patent system is characterized by considerable ambiguity (Lemley and Shapiro, 2005), and patent enforcement is typically a "noisy" process (Jaffe and Lerner, 2011). Particular attention has been paid to the negative impact of patent litigation initiated by NPE, especially those that function as patent trolls, or more precisely as Patent Assertion Entities (PAE), whose business is based on the sole patent enforcement. Further, PAE have a number of advantages in patent litigation compared to manufacturing companies: first, they are often difficult to identify because they hide behind multiple shell companies and names (Hagiu and Yoffie, 2013; Orsatti and Sterzi, 2018); second, they benefit from asymmetries in the potential costs and risks of patent litigation, such as bearing fewer reputational and discovery costs, not being vulnerable to countersuit, and not risking their business if a patent is invalidated (Yeh, 2012). For these reasons, PAE are thought to increase innovation cost, because of the costs of a legal process that can yield rents from producing, innovative firms and of the risk that imperfect courts may rule in the PAE's favor, even if no patent infringement has actually occurred.

The costs associated with PAE litigation have prompted an increasing number of scholars to highlight the negative effects of this type of opportunistic litigation, which is believed to stifle innovation and growth (Bessen and Meurer, 2008; Boldrin and Levine, 2002; Bessen and Meurer, 2013*b*) and therefore requires new legislative efforts to curb the activities of trolls (Curtin, 2014). Chien (2013), for example, presents evidence that PAE have a detrimental impact on entrepreneurial activity. She shows that forty percent of small businesses that received a demand reported a "significant operational impact." For example, the company delayed hiring employees or the achievement of another milestone, changed its product, shifted its business strategy, shut down a business line, or reported a loss in potential investor perception of value. The evidence that PAE litigation is opportunistic and not necessarily indicative of the actual technology covered by the patent is on the spotlight. In particular, PAE target companies that are " flush with cash" and companies that have experienced recent positive cash shocks, even if their profits come from businesses unrelated to the allegedly infringed patents (Cohen et al., 2019).

PAE opportunistic behavior may increase because there are greater rewards for protecting IPR in a certain sector, leading to a shift in the payoff of litigation as a result of changing economic conditions (Kiebzak et al., 2016). However, the number of patent litigations may also increase as a result of a change in incentives in the legal system that makes opportunistic enforcement of patents more lucrative. This is consistent with prior research by Bessen (2005), who view patent litigation as a reflection of the "stakes" in a certain business at a given period.

In this research, we examine the interplay between environmental incentives, proxied by GPP, and the likelihood of litigation of green patents. GPP aims to encourage the local development of new technologies that can facilitate the achievement of environmental sustainability goals. This is because the successful development of green technologies requires navigating a high degree of uncertainty (Mowery et al., 2010). From this perspective, site-specific GPP is considered a direct form of government action to stimulate demand for GTs (Parikka-Alhola, 2008). However, the incentives for new markets and the associated revenues may entice opportunistic behavior by individuals and companies, such as PAE, who seek to profit from them. Consistently, we put forward the following hypothesis:

Hypothesis 2 (H2): *Green patents have a greater probability of being litigated by PAE when there has been green public procurement*

2.3 Research design

In this section, we introduce the data and the variables that we will later use for the empirical analysis. We build our database from a number of sources. First, we collect data on patent litigation in the years 2001-2016. Second, we match these data with other publicly available databases of US patent data (USPTO). Third, we obtain information on patent characteristics and geolocalize patents at the county level; finally, we complete our data with information on geospatial data about green public procurement which vary across different years and different counties.

2.3.1 Data and Variables

Litigation data

We identify litigation cases using the Stanford NPE Litigation Dataset and the Patent Litigation USPTO Dataset. The Stanford NPE Litigation Dataset is the first comprehensive patent litigation dataset that categorizes all US patent lawsuits filed since 2000. It was created by Stanford Law School students under the direction of Mark Lemley and Shawn Miller. The dataset includes nearly 70'000 patent lawsuits filed from 2000 to 2019, covering 62,'195 patents, with more than 80% of cases categorized as including practicing entities or one of eleven types of NPE patent asserters. To identify PAE among the various plaintiffs, the dataset includes a variable called "Asserter Category". Specifically, PAE refer to companies that fall into Category 1 (patents acquired), Category 4 (corporate inheritance), or Category 5 (company formed by a single inventor). In addition, each record contains the identification number of the litigated patent, the number of the related case, the date of filing of the litigation and the venue chosen for the litigation (Miller et al., 2017).

The USPTO Patent Litigation Dataset instead contains data on litigation that occurred between 1963 and 2019, for a total of 81'350 unique district court cases and 45'767 patents. The dataset consists of four distinct subsets linked via the case ID, each containing different information: the first about the case itself, the second about the parties involved, the third about the attorneys involved in the case, and the last about the patents involved in each case. In addition, the team coded the type of case for all of these actions, such as infringement, declaratory judgment, false marking, ownership dispute, malpractice, etc. (Marco et al., 2015); in this context, a small but significant number of cases identify multiple case types (up to three). Therefore, three case type variables (case_type_1, case_type_2, and case_type_3) are included. About 85 percent of the cases are exclusively about patent infringement, and another 8 percent are requests for declaratory judgments. For the purposes of this study, only infringement cases are included in the analysis (case_type=1).

Merging these two datasets yields 184'045 observations associated with 62,195 litigated patents, 10,015 of which are litigated by PAE. This is possible because each patent may be involved in more than one lawsuit over time and may also be litigated in different courts.

Patent data

The second step is to link the litigation information to other relevant patent attributes, such as the assignee of the patent and the year the patent was filed as well as to the quality characteristics. To do this, we use USPTO patent data from Patents View, a platform built on a regularly updated database that links inventors, their organizations, locations, and reports the name(s) of the first assignee(s). Starting from the Assignee table, we can identify patents assigned to NPE by calculating the edit distance (Levenshtein distance) between the assignee name and a list of identified NPE and subsidiaries. Furthermore, for each patent we retrieve other characteristics from USPTO-OECD patent quality database (version 2019).

The third step is to identify green patents among other patents. Patents are defined as green according to two established international classifications, both based on the International Patent Classification (IPC): WIPO IPC Green Inventory (WIPO, 2012) and OECD ENV-TECH (Haščič and Migotto, 2015). We put together both classifications to define our dependent variable Green (Ghisetti and Quatraro, 2017). The variable *Green* is a dummy variable that takes the value of one if the patent is classified as environmentally-related by either the WIPO IPC Green Inventory or the OECD ENV-TECH classification, and zero otherwise. Using this definition of green patent, we identify 197'553 green patents filed with the USPTO between 2001 and 2016 (around 9.2% of the 2'159'065 patents filed during the same period).

Finally, high-tech related patents are selected according to the definition of high-tech patents proposed by Eurostat, using specific subclasses of the International Patent Classification (IPC) defined in the trilateral statistical report of the EPO, the JPO and the USPTO. The following (macro) technical fields are defined as high technology: Computer and automated business equipment; Microorganism and genetic engineering; Aviation; Communication. Using this classification, we identify 1'066'794 high-tech patents.

In our dataset, NPE are assignees of 51'440 patents, among which 41'901 are ICT related patents. Moreover, ICT patents represent 81.5% of the patents litigated by PAE.

For what concerns the geographical dimension, we assign a patent to a US county by means of information contained in assignees' addresses. To do that, we use a combination of the location dataset developed in 2019 by Prof. Gaétan de Rassenfosse which provides geographic coordinates for assignee locations in 18.8 million patent documents spanning over more than 30 years (De Rassenfosse et al., 2019), the raw location dataset of Patents View and the USPTO assignee dataset[1]. The geocoded data are

[1] Furthermore, when even using these sources we could not find the necessary detail to geolocalize the patent at county level, we used two different strategies: 1-in the first place, for those patents missing zip codes, we recovered them using a public API (`https://geo.fcc.gov/api/census/area`) to get them using their latitude and longitude 2-eventually, for those still lacking county information despite having a zip code, we used a crawler to scrape zip codes from unitedstateszipcodes.org and to get in that way the county of the remaining patents. The results and the code for the crawler are public and available on this github (`https://github.com/graziasveva/Patent_location_US`)

further allocated to the corresponding countries, regions and cities. In our cases, only patents filed at the American patent office and whose assignees addresses lie within US are selected.

Green public procurement data

We obtained data on environmental procurement spending from public information sources. For this analysis, we use procurement data retrieved from USAspending.gov. The information retrieved is for all registered federal contracts, from which we determine the location of funding provision (county level) where the contract is executed and the amount of funds used for each contract. Then the county-level information is grouped by year.

Furthermore, we use the Product and Service Codes Manual (PSC, August 2015 edition) to identify procured green contracts. The PSC Manual gives criteria to classify products, services, and R&D purchased by the federal government for each contract action reported in the Federal Procurement Data System (FPDS). Since a contract may contain multiple products/services with and without environmental attributes, the PSC data element code was selected based on the predominant product or service purchased. Environmental purchases are defined based on the criterion that the products or services have less or fewer impacts on human health and the environment compared to competing products or services that serve the same purpose. Environmental attributes refer to areas such as 'energy efficient', 'bio-based', or 'environmentally preferable'. The FPDS acknowledges that determining environmental attributes is challenging because, for example, the purchase of services is usually predominantly labour-intensive and therefore does not necessarily clearly entail environmental attributes. Nevertheless, environmental attributes can be important in achieving the goals of the contract (Orsatti, Perruchas, Consoli and Quatraro, 2020).

Descriptive analysis

Table 2.1 reports the main descriptive statistics for the variables used in the analysis. Figure 2.1 and Figure 2.2 provide a visual summary of the geographic distribution of the two dimensions of interest: the number of green patents filed in each county from 2001 to 2016 and the level of GPP in the same geographic areas with a one-year lag. In both figures, each county is outlined in black, while the color is shaded according to the quintile rank distribution and becomes progressively darker for higher quintiles. In addition, the quintiles are weighted based on population density. Figure 2.1 shows that most green patents are awarded in and around coastal metropolitan areas, such as Seattle, San Francisco, Los Angeles in the West, and Miami, Orlando, and New York on the East Coast. There are also some interior areas relevant to green patents, such as the Denver area in Colorado and the Boston area in Massachusetts.

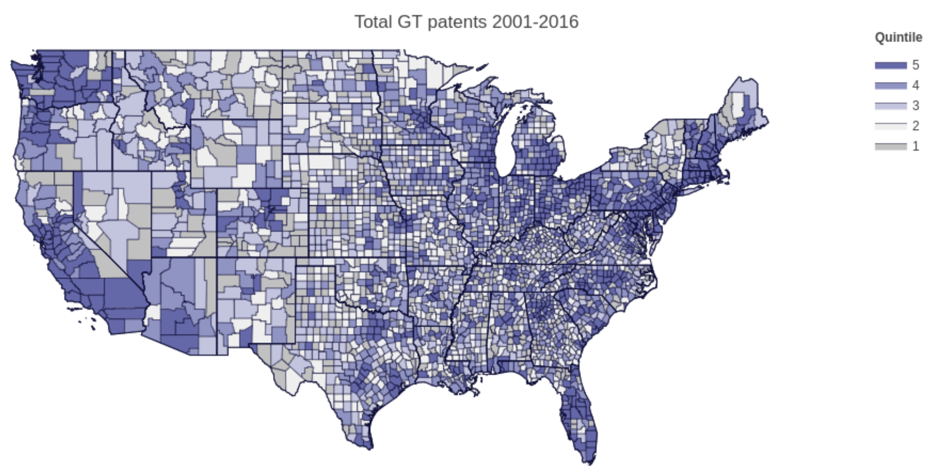
Figure 2.2 instead presents the geographic quintile distribution of total county-level GPP spending levels from 2000-2015. Looking at both maps together, it is clear that there is some overlap between

Table 2.1: Descriptive statistics (2001-2016)

VARIABLE	DESCRIPTION	OBSERVATIONS	MEAN	SD	MIN	MAX
litigation	litigated patent (dummy)	2'159'065	.0091	0.095	0	1
PAE*	PAE litigated patent (dummy)	19'785	.208	0.406	0	1
green	green patent (dummy)	2'159'065	.091	0.288	0	1
GPP dummy	GPP in county j at time t-1 (dummy)	2'159'065	.537	0.498	0	1
GPP log	log of the GPP in county j at time t-1	2'159'065	8.967	8.852	0	28.436
patent scope	no. of distinct IPC codes	2'159'065	1.988	1.241	1	30
bwd cits	no. of backward cits	2'159'065	29.348	35.877	0	315
npl cits	no. of non-patent literature cits	2'159'065	7.767	17.495	0	102
claims	no. of distinct claims	2'159'065	19.05	11.847	1	255
fwd cits5	no. of cits in 5 years from the filing date	2'159'065	14.86	48.730	0	1'964
NPE_owned	patent owned by an NPE (dummy)	2'159'065	.023	0.152	0	1
ICT	patent belongs to one of the first 8 WIPO tech. classes (dummy)	2'159'065	.494	0.499	0	1
Texas*	litigation filed in the court of Texas (dummy)	19.'785	.116	0.320	0	1
age*	age of the patent at time of litigation	19.'785	6.32	3.47	0	15

Notes: Unit of observation:patent. Filing years 2001-2016. *Information about PAE litigation, age and the venue of litigation is available only for litigated patents.

Figure 2.1: Green patents assigned within the US

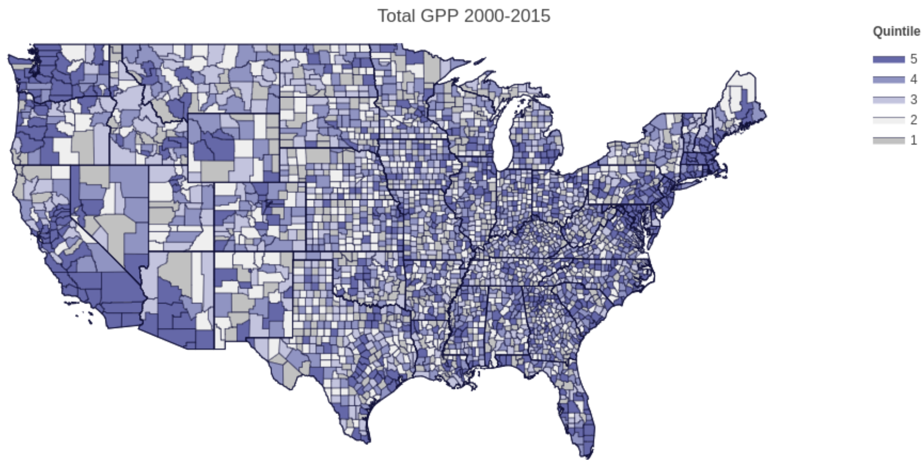


Notes: Geographic quintile distribution of green patents, 2001–2016. Quintiles are weighted by population density.

the distribution of green patents and the distribution of GPP. However, both maps show a great deal of heterogeneity across states and counties.

In addition, Table 2.2 provides further detail on patent asserters belonging to either producing companies or PAE; in particular, it shows the top 10 patent asserters by the number of litigated patents between 2003 and 2016. The topics of non-PAE patent litigation are more diverse and cover a wide range of technological classes, such as processes for making chemical compounds (C12P), batteries for converting chemical energy into electrical energy (H01M), and systems for regulating electrical or magnetic variables (G05F). Consequently, patent asserters in this category address a variety of technological products related to green technologies. Butamax (TM) Advanced Biofuels, for example, is a joint venture that combines expertise in fuels with industrial biotechnology. Further, Metco Battery Technologies is a global company

Figure 2.2: Green public procurement



Notes: Geographic quintile distribution of GPP total expenditures, 2000–2015. Quintiles are weighted by population density.

Table 2.2: TOP 10 patent asserters by category

No.	Product companies	IPC/CPC most litigated class	Patent Assertion Entities	IPC/CPC most litigated class
1	Butamax (TM) Advanced Biofuels LLC	C12P (fermentation to create chemical compunds)	Double Rock Corporation	G06Q (data processing system)
2	Metco Battery Technologies LLC	H01M (batteries, electrical energy)	Global Touch Solutions, LLC	G06C (digital computing)
3	Milwaukee Electric Tool Corporation	H01M (batteries, electrical energy)	Island Intellectual Property, LLC	G06Q (data processing system)
4	Philips Solid-State Lighting Solutions, Inc.	G05F (systems to regulate electric variables)	Skky, LLC	H04L (trasmission of digital information)
5	Koninklijke Philips Electronics N.V.	G05F (systems to regulate electric variables)	Sipco, LLC	G08G (traffic control systems)
6	Server Technology, Inc.	G01R (measuring electric variables)	D Three Enterprises, LLC	E04D (roof covering)
7	AC (Macao Commercial Offshore) Limited	H01M (batteries, electrical energy)	Secured Mail Solutions, LLC	G06Q (data processing system)
8	Techtronic Industries Co., Ltd.	H01M (batteries, electrical energy)	Solannex, Inc.	H01L (semiconductor devices)
9	Nortek Air Solutions, LLC	F01D (steam turbines)	The Abell Foundation, Inc.	B60K (propulsion unit in vehicles)
10	SynQor, Inc.	G05F (systems to regulate electric variables)	Customedia Technologies L.L.C.	H04N (transmission of signal)

Notes: This table presents the top 10 patent asserters in terms of number of litigated patents from 2003 to 2016. Duplicates (patent litigated more than once) are not considered to make the rank.

specializing in energy solutions. On the other hand, if the green patent litigations involve PAE, most of them relate to high-tech areas such as data processing systems (G06Q), digital traffic control devices (G08C), and so on. Thus, the main business of the companies defined as PAE seems to be far from green technologies. Double Rock Corporation, for example, is described as a leader in cash management and financial technology, without any reference to green technology. The same is true for Global Tech Solution, which provides IT support such as technical help desk support, computer support, and consulting services to other companies.

2.3.2 Empirical strategy

In our first econometric exercise, the response variable (*Litigation*) is thus a dichotomous variable whereby litigated patents are designated as "1" while non-litigated ones are coded with "0". Hence a Logit model approach is deemed appropriate here:

$$Prob(Litigation_i = 1) = F(X_i) = \frac{e^{\beta' X_i}}{1 + e^{\beta' X_i}}$$

In other words, the probability of a litigation is assumed to be a function of a vector of explanatory variables, X. Our main variable of interest is the dummy variable which identifies green patents (*Green*), as well as the variable which represents the level of expenditures for green public procurement performed in each county j at time *t-1*. In the different models it is represented either by the log transformed green expenditure (GPP_log) or by a dummy which takes value 1 when the GPP in county j at time t-1 is greater than 0, 0 otherwise (GPP_dummy). Then, we include respectively the interaction between the green dummy and either the GPP_dummy variable (Model 1) or the GPP_log (in the latter case the interaction variable is called $green\_GPP\_log$) (Model 2), adding then controls for some relevant patent characteristics represented by the vector X', such as the number of forward citations (fwd_cits5), number of claims (*claims*), patent scope (*patent_scope*), backward citations (*bwd_cits*) and non-patent literature citations (*npl_cits*). We add year dummies to control for changes across years that affect the probability of litigate. Standard errors are clustered at patent level.

Model 2.1 can be mathematically described as follows:

$$Prob(Litigation_{j,t} = 1) = \beta_0 + \beta_1 Green_{j,t} + \beta_2 GPP\_dummy_{j,t-1} + \beta_3 Green * GPP\_dummy + \beta_4 X^{'}_{j,t} + \epsilon_{j,t} \tag{2.1}$$

In Model 2.2, we substitute the GPP_dummy with the GPP_log variable and the interaction term $green\_GPP\_log$:

$$Prob(Litigation_{j,t} = 1) = \beta_0 + \beta_1 Green_{j,t} + \beta_2 GPP\_log_{j,t-1} + \beta_3 green\_GPP\_log_{j,t} + \beta_4 X^{'}_{j,t} + \epsilon_{j,t} \tag{2.2}$$

In Model 2.3 we include the NPE_owned dummy which is a variable that takes value 1 when a patent has been owned by an NPE in its history, 0 otherwise. We interact NPE_owned with the $Green$ and GPP_dummy:

$$\begin{aligned} Prob(Litigation_{j,t} = 1) = \beta_0 + \beta_1 Green_{j,t} + \beta_2 GPP\_dummy_{j,t-1} + \beta_3 NPE\_owned_{j,t} \\ + \beta_4 Green * GPP\_dummy * NPE\_owned + \beta_5 X^{'}_{j,t} + \epsilon_{j,t} \end{aligned} \tag{2.3}$$

Eventually, in Model 2.4, we add as independent variable the ICT dummy which takes value 1 when the patent is high-tech, 0 otherwise. Then we add an interaction term between the $Green$ and ICT variables. Model 2.4 looks as follows:

$$\begin{aligned} Prob(Litigation_{j,t} = 1) = \beta_0 + \beta_1 Green_{j,t} + \beta_2 GPP\_dummy_{j,t-1} + \beta_3 ICT_{j,t} \\ + \beta_4 Green * ICT + \beta_5 X^{'}_{j,t} + \epsilon_{j,t} \end{aligned} \tag{2.4}$$

In our second econometric exercise we investigate the relationship between the variables above mentioned and a sub-sample of the litigated patents, the patents litigated by PAE. The dependent variable PAE takes value 1 when the patent has been litigated at least once by a PAE, 0 when the patent has been litigated but not by a PAE. The first two models mirror the equations from 2.1-2.2 and add the ICT variable, which is demonstrated to be especially relevant in PAE related litigation (US-FTC, 2016; Sterzi et al., 2021*b*). Using PAE as dependent variable Model 2.5 and Model 2.6, are of the following form:

$$\begin{aligned} Prob(PAE_{j,t} = 1) = \beta_0 + \beta_1 Green_{j,t} + \beta_2 GPP\_dummy_{j,t-1} + \beta_3 \\ Green * GPP\_dummy + \beta_4 ICT_{j,t} + \beta_5 X^{'}_{j,t} + \epsilon_{j,t} \end{aligned} \tag{2.5}$$

$$\begin{aligned} Prob(PAE_{j,t} = 1) = \beta_0 + \beta_1 Green_{j,t} + \beta_2 GPP\_log_{j,t-1} + \beta_3 green\_GPP\_log_{j,t-1} \\ + \beta_4 ICT_{j,t} + \beta_5 X^{'}_{j,t} + \epsilon_{j,t} \end{aligned} \tag{2.6}$$

We decided to add to Model 2.7 and Model 2.8 two other variables which are usually related with litigation involving PAE. The first one is the *age* of the patent, which has been calculated subtracting the year of patent application to the first litigation year (or simply to the litigation year if the patent has been litigated one time), included in Model 2.7. In most of the cases, PAE are said to litigate older patents (Fischer and Henkel, 2012; Love, 2013; Feng and Jaravel, 2016). In Model 2.8, we add another control related to PAE patent litigation, the $Texas$ dummy, which takes value 1 when the patent is litigated in the Texas District (a place which is well known to be favorable to PAE as argued by Allison et al. (2013) and Liang (2010)) and 0 otherwise. Adding these two variables, Model 2.7 and Model 2.8 look respectively as follows:

$$\begin{aligned} Prob(PAE_{j,t} = 1) = \beta_0 + \beta_1 Green_{j,t} + \beta_2 GPP\_log_{j,t-1} + \beta_3 green\_GPP\_log_{j,t-1} \\ + \beta_4 ICT_{j,t} + \beta_5 Age_{j,t} + \beta_6 X^{'}_{j,t} + \epsilon_{j,t} \end{aligned} \tag{2.7}$$

$$\begin{aligned} Prob(PAE_{j,t} = 1) = \beta_0 + \beta_1 Green_{j,t} + \beta_2 GPP\_log_{j,t-1} + \beta_3 green\_GPP\_log_{j,t-1} + \\ \beta_4 ICT_{j,t} + \beta_5 Age_{j,t} + \beta_6 Texas_{j,t} + \beta_7 X^{'}_{j,t} + \epsilon_{j,t} \end{aligned} \tag{2.8}$$

Finally, Model 2.9 mirrors the specifications of Model 2.4, studying the interaction term between the $Green$ and ICT variables in the context of PAE litigation:

$$
\begin{aligned}
Prob(PAE_{j,t}=1) = \beta_0 + \beta_1 Green_{j,t} + \beta_2 GPP\_dummy_{j,t-1} + \beta_3 ICT_{j,t} \\
+\beta 4 Green * ICT + \beta_5 Age_{j,t} + \beta_6 Texas_{j,t} + \beta_7 X^{'}_{j,t} + \epsilon_{j,t}
\end{aligned}
\tag{2.9}
$$

2.4 Results

2.4.1 Litigation involving green patents

Table 2.3 reports the results related to all the litigated patents in our sample. In all the models presented in this table the odds of being litigated for green patents are lower. This is consistent with the literature stressing that complex technologies are more difficult to litigate, hence green patents are not a primary choice when considering the most suitable patents for litigation (Lanjouw and Schankerman, 2001). Moreover, all the quality controls (*patent_scope*, *bwd_cits*, *npl_cits*, *claims* and fwd_cits5) variables are significant and are linked to a higher probability of litigation, as it is argued by those scholars who say that high quality is an important determinant in the litigation decision (Reitzig et al., 2007; Lanjouw and Schankerman, 2001; Bessen, 2005). For instance, a one unit increase in fwd_cits5, leads to a 0.2% increase in the odds of being litigated Model 2.1 (Column I) and Model 2.2 (Column II).

In Model 2.1 we have an interaction term between the *Green* and GPP_dummy variables, thus the interpretation of the coefficients is different from the models with no interaction. In particular, the odds of being litigated for a green patent are 0.667 times (34% lower) those of a non-green patent in the case in which there is no green public procurement in the county (the reference case for the GPP_dummy is 0). At the same time, the coefficient for the GPP_dummy indicates that the odds for a non-green patent filed where GPP is greater than 0 (GPP_dummy=1) are 0.870 times (13% lower) the odds of a non-green patent filed where there is no green public procurement. However, the interaction term itself is not significant.

In Model 2.2 (Column II) the coefficients of the *Green* and GPP_log variables are both significant and negative. Considering the interaction term is not significant, we can say that green patents have 34% lower odds of being litigated compared to non-green patents, while an additional unit of GPP_log is associated with a decrease of 0.86% in the odds of being litigated.

Model 2.3 (Column III) confirms the previous results, adding that the odds of being litigated for a green patent, considering all the other terms at their reference value, are 0.666 times (33,5% lower) the odds of a non-green patent. The NPE_owned coefficient is positive and significant indicating that NPE owned patents are at higher risk of being litigated. When we consider the odds of being litigated if the patent is owned by an NPE, for green patents the odds are 1.38 times (38 % higher) the odds of non-green patents that are also NPE-owned . Furthermore, when the patent is owned by an NPE, the odds of being litigated for a patent filed where the GPP is greater than 0 are 1.12 times higher the odds of a patent filed where the GPP 0.

These results highlight the importance of NPE strategies in the litigation landscape and in Section 2.4.2 we further explore this concept with more econometric analyses.

Model 2.4 (Column IV) eventually analyzes the interplay between green and ICT patents; the odds of being litigated for green patents that are not ICT are 0.538 times those of non-green patents that are not ICT. The odds of being litigated for ICT patents that are not green instead are 0.747 times the odds of non-ICT patents. The interaction term in this model can be interpreted as follows: the odds of being litigated for a green ICT patent are 0.86 times the odds of a non-green ICT patent.

2.4.2 PAE litigation involving green patents

Table 2.4 proposes similar models to Section 2.4.1 but focusing on patents litigated by PAE among all the litigated patents. Interestingly, in this case quality variables are mostly not significant except for *claims* and *patent_scope* which are positive and significant in all the models of the Table. Some scholars explain this fact arguing that patent litigated specifically by PAE are of lower quality (Chien, 2011; Feng and Jaravel, 2020). In addition Cohen et al. (2019) state that PAE usually litigate wordier patents with more claims.

The variable ICT is significant and positive as well, consistently with the literature underlining that high-tech patents as more often involved into PAE litigation than other patents (Bessen and Meurer, 2008; Orsatti and Sterzi, 2018; Sterzi, 2021). Progressively from Model 2.7 onwards we add controls for patent age (age) and for litigation filed in Texas, famously home to opportunistic litigation ($Texas$).

In Model 2.5 (Column V), we show that the odds of being litigated by a PAE for a green patent filed where there is no green public procurement are 0.756 times those of non-green patents filed where there is no public procurement. In fact, the odds of being litigated for a green patent filed where there is green public procurement are 6% higher than the odds for a non-green patent filed where there is green public procurement. Model 2.6 to 2.8 (Column VI, VII, VIII) confirm this positive and significant joint impact of GPP and green patents on the probability of litigation. In particular, Column VI shows that one unit increase in GPP_log when the patent is not green decrease the odds of litigation by 0.58 %. On the other hand, for a green patent, a one unit increase in GPP_log yields a change in the odds of 2.13%.

In Model 2.9 (Column IX), the odds of being litigated for a green patent that is not ICT are 1.73 times higher of those of non-green patent that is not ICT. At the same time, the odds for an ICT patent that is not green are 4.65 times higher the odds for a non-ICT patent that is not green. Finally, the odds of being litigated for a green and ICT patent are 0.65 times the odds of a non-green ICT patent.

Table 2.3: Main results. LOGIT regression (all litigated patents)

Y= litigated patent	I Odds ratio	II Odds ratio	III Odds ratio	IV Odds ratio
1.green	0.6676*** (-7.13)		0.6660*** (-6.82)	0.5385*** (-9.98)
green		0.6631*** (-7.35)		
1.GPP_dummy	0.8701*** (-5.95)		0.8391*** (-7.20)	
1.green1.GPP_dummy	0.9764 (-0.30)		0.9477 (-0.64)	
GPP_dummy				0.8951*** (-4.88)
patent_scope	1.0262*** (3.82)	1.0263*** (3.83)	1.0263*** (3.75)	1.0125* (1.74)
bwd_cits	1.0056*** (25.30)	1.0056*** (25.29)	1.0055*** (24.28)	1.0054*** (24.32)
npl_cits	1.0089*** (19.04)	1.0089*** (19.05)	1.0085*** (17.81)	1.0088*** (19.10)
claims	1.0134*** (26.35)	1.0134*** (26.37)	1.0129*** (24.77)	1.0137*** (27.33)
fwd_cits5	1.0020*** (28.99)	1.0020*** (28.94)	1.0019*** (26.72)	1.0022*** (30.37)
GPP_log		0.9914*** (-5.78)		
green_GPP_log		0.9996 (-0.10)		
1.NPE_owned			3.3498*** (29.05)	
1.green1.NPE_owned			2.0723*** (3.80)	
1.GPP_dummy1.NPE_owned			1.3361*** (4.45)	
1.green1.GPP_dummy1.NPE_owned			1.0480 (0.18)	
1.ICT				0.7469*** (-13.35)
1.green1.ICT				1.5986*** (5.69)
Observations	2159065	2159065	2159065	2159065
Pseudo R2	0.0770	0.0770	0.0884	0.0787
Year Dummies	YES	YES	YES	YES

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: Unit of observation: litigated patent. Litigation years: 2003-2016. The results are expressed in odds ratios (OR); t statistics in parentheses.

2.5 Discussion and conclusion

In recent decades, we have seen an increasing attention to environmental issues, and in the transition to a greener economy, the development and adoption of eco-innovations is considered crucial. Green technologies aim to solve complex problems, such as reducing environmental impact, and to this end they are designed as high quality, sophisticated technologies (del Río González, 2009; Barbieri et al., 2020). The adoption of green technologies is encouraged by a stringent regulatory framework. The stricter the

Table 2.4: Main results. LOGIT regressions (PAE litigated patents)

Y= PAE litigated patent	V Odds ratio	VI Odds ratio	VII Odds ratio	VIII Odds ratio	IX Odds ratio
1.green	0.7599* (-1.77)				1.7366*** (3.03)
green		0.7217** (-2.05)	0.7267** (-1.97)	0.7558* (-1.75)	
1.GPP_dummy	0.8653*** (-2.69)				
1.green1.GPP_dummy	1.3912* (1.67)				
GPP_dummy					0.8837** (-2.24)
ICT	5.5241*** (29.25)	5.5063*** (29.20)	5.1551*** (27.82)	4.4115*** (24.62)	
patent_scope	1.0595*** (3.06)	1.0595*** (3.06)	1.0605*** (3.04)	1.0661*** (3.33)	1.0640*** (3.24)
bwd_cits	1.0004 (0.52)	1.0004 (0.52)	1.0010 (1.46)	1.0011 (1.43)	1.0011 (1.54)
npl_cits	1.0004 (0.37)	1.0004 (0.37)	0.9999 (-0.13)	1.0001 (0.06)	1.0001 (0.06)
claims	1.0033*** (2.79)	1.0033*** (2.79)	1.0027** (2.06)	1.0026* (1.90)	1.0025* (1.85)
fwd_cits5	1.0003 (1.30)	1.0003 (1.27)	1.0004 (1.62)	1.0002 (1.04)	1.0002 (1.03)
GPP_log		0.9942* (-1.69)	0.9946 (-1.59)	0.9932* (-1.91)	
green_GPP_log		1.0273** (2.30)	1.0262** (2.18)	1.0270** (2.20)	
age			1.1999*** (15.96)	1.1780*** (14.27)	1.1785*** (14.27)
Texas				4.0900*** (19.74)	4.0720*** (19.64)
1.ICT					4.6546*** (24.59)
1.green1.ICT					0.3738*** (-4.50)
Observations	19785	19785	19785	19785	19785
Pseudo R2	0.1158	0.1158	0.1482	0.1857	0.1869
Year Dummies	YES	YES	YES	YES	YES

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: Unit of observation: litigated patent. Litigation years: 2003-2016. The results are expressed in odds ratios (OR); t statistics in parentheses.

regulatory framework, the greater the incentives for companies to comply with it (Orsatti, Perruchas, Consoli and Quatraro, 2020). This system is particularly beneficial if it favors demand-pull mechanisms to create new market niches and thus encourage the diffusion of green technologies (Rennings, 2000). However, this could attract different actors that want to benefit from the profit prospects opened by the implementation of environmental policies. In this paper, we have specifically examined the interplay between the strategies of different actors in patent litigation and the regulation induced prospects in the market for green technologies; in this sense, our contribution is threefold. First, we are among the first to investigate patent litigation in the green technological domain. In particular, we have examined whether green innovations that are perceived to have different characteristics, typically involving higher quality and more sophisticated knowledge, are at greater risk of becoming targets of patent litigation. Relevant scholars studying litigation emphasize that there is a strong and positive relationship between the likelihood of litigation and the quality characteristics of a patent (Allison et al., 2003; Bessen, 2005; Lanjouw and Schankerman, 2001). However, our result seems to point in a different direction, showing that green patents do not have a higher risk of being litigated. One possible explanation could be related to the greater breadth of green patents (Barbieri et al., 2020; Dechezleprêtre et al., 2013): green technologies draw on slightly more diversified knowledge fields and, in particular, combine a much bigger number of technological components; on the point, Lanjouw and Schankerman (2001) say that although patent breadth, i.e., the number of different disciplines related to a given patent, can be considered a proxy for its market value (Lerner, 1994), an increase in patent breadth is associated with a lower probability of litigation. This because, as mentioned by the aforementioned authors, it is difficult for the patent holder to identify infringement when the invention involves different technological fields.

Second, when we examine litigation involving NPE the situation is different, as green patents owned by an NPE have a greater probability of being litigated compared to non-green patents. Third, when we check whether stringent environmental regulations might be a possible trigger for green patent litigation, the situation is different for litigation vs. PAE litigation. In the first case, there seems to be no significant relationship between higher GPP and higher risk of patent litigation for green patents; in the second case in fact, there is a positive and significant effect of GPP on the likelihood of a green patent being litigated. This could be interpreted as an indication of a possible strategic behavior of patent trolls who prefer to litigate where there are stricter environmental regulations. This is corroborated by the fact that lead plaintiffs identified as PAE in our data do not seem to have a clear connection to green technologies and therefore may be participating in this type of litigation, not to defend their property rights, but because of greater profit prospects. These results shed light on the "dark side" of environmental stringency: while stricter environmental regulations are intended to encourage the adoption and market for green technologies, they may actually favor strategic behaviors by those who own these technologies. The policy implications of these findings call for a closer look at these "side effects" of stricter environmental regulations, which in this case could increase the cost of adoption of green technologies and create a profitable market for patent trolls.

Chapter 3

Technological diversity to address complex challenges: the contribution of American universities to SDGs

3.1 Introduction

The Sustainable Development Goals (SDGs) aim to address the complex challenges of our century through 167 interlinked and interdisciplinary targets; progress on one goal depends on and influences other goals. For example, developments in agriculture towards zero hunger (SDG 2) depend on affordable and clean energy (SDG 7), while also reducing inequalities (SDG 5 and SDG 10) and protecting life on land (SDG 15) and in water (SDG 14). The use of technology and innovation is crucial to achieving the goals of the SDGs, as it fosters the knowledge economy and the resulting creation of inventions that can help decouple economic growth from the risk of environmental and social crises and improve living conditions in areas such as the environment, energy, medicine and transport (Blohmke, 2014; Walz et al., 2017). In this sense, technological development can unleash its potential for systemic change while having a positive impact on the environment and society (Deuten, 2003). The importance of science, technology and innovation to reach the SDGs has been stressed in the UN official documents where they are intended to boost the capacity of countries to change the current trajectory and accelerate progress toward a sustainable future (UN, 2020). However, there is no direct reference to IP in the goals and targets of the 2030 Agenda, with the exception of paragraph 3.b of Goal 3, nor any IP-related indicators in the current Global Indicator Framework. Therefore, in this research we contribute to reduce this gap, proposing a novel methodologies to map patents to the SDGs, adding evidence about the relationship between IP and SDGs.

Although the literature linking interdisciplinarity to the SDGs is extensive, less attention has been dedicated to the characteristics of innovation in the context of the SDGs (van der Waal et al., 2021;

Hajikhani and Suominen, 2021) and recent studies do not consider potentially relevant stakeholders such as universities, which are expected to contribute to the achievement of the SDGs through a mix of education, research and innovation (Owens, 2017; Sánchez-Carracedo et al., 2021; Kopnina, 2020). Further, a popular view is that exploiting a single domain promotes "one way of thinking, damping creativity, while combining knowledge from diverse and distant domain leads to more breakthrough innovation." (Hargadon and Sutton, 1997; Ahuja and Morris Lampert, 2001). This has been proven to be the case for green technologies, which are characterized by intrinsic complexity and therefore result from the integration of different and heterogeneous technologies and knowledge sources (Quatraro and Scandura, 2019; Barbieri et al., 2020; Fusillo et al., 2020). Therefore, this research builds on the green innovation literature and explores the characteristics of SDGs-related innovation, which includes not only green innovation but also the so-called "blue" innovation which relates to unmet sustainable development needs, such as reducing poverty and hunger, promoting health and well-being, education, biodiversity, water and sanitation (van der Waal et al., 2021). In particular, we investigate if SDGs-related innovation is more diverse than its counterpart, as this information might add on the debate about the best policy intervention to foster the development this kind of technologies.

In addition, considering that universities and research centers around the world have made significant progress in establishing collaborative, interdisciplinary initiatives in sustainability science thanks to their more diverse knowledge and skills base (De Marchi, 2012; De Marchi and Grandinetti, 2013) , this research also examines whether universities are able to exploit their favorable position to produce more diverse innovation when it is linked to the SDGs.

Thus, the contribution of this paper to the literature is threefold. First, exploiting the textual part of patents (title, abstract and claims), we develop a novel methodology for tagging SDGs-related patents through an unsupervised natural language processing (NLP) approach; starting from a pre-validated list of keywords, we create a keywords' dictionary for each SDG based on patent text. To do that, we combine the TF-IDF (Term Frequency-Inverse Document Frequency) method with a vectorial representation of patent text. Thanks to the enriched vocabulary, we manage to better identify those patents that were missed in the initial matching due to the peculiarities of the legal jargon characterizing patents in general, but especially claims (Bonino et al., 2010; Tseng et al., 2007). This is, to our knowledge, one of the first attempts to create a proxy measure to analyse the progress towards the SDGs in the innovation system.

Second, we are among the first to analyse innovation related to each SDG, providing original descriptive evidence about that; then we empirically compare the diversity of American SDGs and non-SDGs patents across the main technological classes.

Third, we provide evidence about American universities SDGs patent portfolios composition and investigate whether and for which specific SDGs, there is a diversity premium.

Our results show that, overall, patents related to the SDGs are on the rise, but the trend is more pronounced for patents owned by universities, highlighting that universities are increasingly aware of their role in the sustainability journey; however, most of the production of university patents related to the SDGs seems to revolve around SDG 3 (good health and well-being). Overall, the rise of SDGs patents seem to be led not only by green technologies, but mostly by high technologies, a relationship which has been underestimated by the literature (Vinuesa et al., 2020; Kostoska and Kocarev, 2019). Furthermore, we prove that SDGs related patents are more diverse than non-SDGs patents across most of the main technological fields. Although the role of universities in providing an interdisciplinary perspective to the SDGs is highlighted in the literature, university patents only have a diversity premium for few of the SDGs (namely SDG 2, SDG 3, and SDG 15). This may suggest that interdisciplinarity is seen as a valuable resource in universities, but paradoxically it is more difficult to achieve in sustainable innovation.

The rest of the paper is organized as follows: Section 2 presents the theoretical background and the research hypotheses, Section 3 presents the Research Design, Section 4 shows the main results and Section 5 concludes.

3.2 Theoretical Background and hypotheses development

3.2.1 Intellectual property and the SDGs

Intellectual property (IP) is a critical incentive for innovation and creativity, which are key to achieving the Sustainable Development Goals (Walz et al., 2017; Cordova and Celone, 2019). Only through human ingenuity it is possible to develop new solutions not only to promote economic growth, but also to eradicate poverty, increase agricultural sustainability and ensure food security, combat disease, improve education and equality, protect the environment, and accelerate the transition to a low-carbon economy to combat climate change and preserve biodiversity (Rimmer, 2018). From this perspective, innovation and creativity are not goals in themselves; they are methods and tools for innovative solutions to development problems and, because they are at the core of the system, have an impact on a number of SDGs. As such, technologies are deemed to directly impact SDG 2 (zero hunger) (Blakeney, 2009; Oguamanam, 2006), SDG 3 (good health and well-being) (Abbott, 2002), SDG 6 (clean water and sanitation), SDG 8 (decent work and economic growth) and SDG 9 (industry, innovation and infrastructure) (WIPO, 2019). In addition, increasing the share of environmental oriented technologies is essential to achieve SDG 7 (affordable and clean energy), SDG 11 (sustainable cities and communities), SDG 12 (responsible production and consumption), SDG 13 (climate action), SDG 14 (life under below water) and SDG 15 (life on land) (Henry and Stiglitz, 2010; Rimmer, 2014). Further, the so-called "blue" technologies, namely those aiming at "improving conditions" (van der Waal et al., 2021) might help in achieving SDG 1 (no poverty) (Idris, 2003), SDG 4 (quality education), SDG 5 (gender equality) and SDG 10 (reduce inequalities) (WIPO, 2019). In a perfect world, a sustainable development agenda for IP would include a "universal call to action" to ensure that the IP system helps address the sustainability related issues (Bannerman, 2020). However, it should be noted that there is no direct reference to IP in the goals and targets of the 2030 Agenda, with the exception of paragraph 3.b of Goal 3, which mentions IP rights in relation to flexibilities to protect public health. In addition, there are no IP-related indicators in the current Global Indicator Framework adopted by the UN Statistical Commission, the UN Economic and Social Council, and the UN General Assembly in 2017. On the one hand, the World Intellectual Property Organization (WIPO) has considered relatively few of its activities to contribute directly to the SDGs, limiting its contribution to explicitly acknowledge the role of IP for SDG 9 and proposing an accurate classification of green technologies (the WIPO *Green Inventory*) that can be used to spot environmental related patents which might be consistent with some SDGs objectives such as those of SDG 6, SDG 7, SDG 13, SDG 14, and SDG 15 (Walz et al., 2017; Guo et al., 2020). Consistently, several scholars emphasize the role of green technologies in fulfilling the 2030 Agenda. For example, the study by Walz et al. (2017) examines the dynamics of green energy and resource efficiency innovation, looking at the position of northern and emerging economies. Instead, Guo et al. (2020) examine the characteristics of sustainable development in the context of green technology, using the indicators of the Sustainable Development Goals Index (SGDI) in its environmental component. On the other hand, researchers urge a more in-depth examination of innovation and the SDGs, particularly in relation to social issues. This call has been echoed by scholars, such as van der Waal et al. (2021) and Hajikhani and Suominen (2021), and practitioners; for instance, IP specialist consultancy *Lex Machina* recently published a study describing an additional feature of the Lexis Nexis database to map patents connected to the SDGs. They emphasize that mapping patents to the Sustainable Development Goals allows companies to objectively measure progress, understand their portfolio and that of their competitors, identify licensing and M&A targets, and assess risks and opportunities in the context of sustainable development. From a policy perspective, SDGs patent mapping could support strategic decision-making for sustainable investment, as well as identify gaps in sustainable technology development and the most effective innovations. This can lead to improved sustainability-related decision making and reporting, as well as influence R&D investments and support investment plans. Therefore, our first research contribution consists in an original attempt of mapping of patents related to each SDG, aiming to identify the key enabling technologies, therefore expanding the literature with new evidence about the relationship IP and the SDGs. The methodological section (Section 3.3.2) provides details about the methodology through which we accomplish the task.

3.2.2 Technological diversity and the SDGs

Many authors believe that inventions result from the combination of existing ideas and devices (Weitzman, 1998; Arthur, 2007). Recombinant inventions are frequently referred to be breakthrough because, in contrast to incremental innovation, merging information from various and distant disciplines nurtures creativity and promotes innovative ideas that are more likely to result in valuable inventions (Ahuja and Morris Lampert, 2001; Audia and Goncalo, 2007; Zhu et al., 2022). This type of innovation is consistent with that required to address sustainability-related challenges (Lam et al., 2014; Jones et al., 2010), where an interdisciplinary approach is required to adequately map multi-layered, complex issues, and without such an approach, necessary solutions risk not be identified. Complex systems, such as acid rain or rapid population expansion, are multifaceted, and standard disciplinary techniques are severely limited in their ability to provide a full view of such phenomena by examining them from the perspective of a single discipline. Newell et al. (2001) suggests that a complete understanding of the interrelationships and dynamics between the various components of these complex events can be achieved by drawing on and integrating multiple perspectives. To this end, it is necessary to combine the efforts of experts from multiple disciplines to address the complex socio-ecological problems of our time (Morse et al., 2007).

Interdisciplinarity has already been proven as an effective means of addressing complex challenges in the innovation system, such as those posed by climate change and environmental degradation. In particular, the link between diversified knowledge sources and green innovation has been analyzed through the framework of recombinant technologies and recombinant competences (Fusillo et al., 2020; Orsatti, Quatraro and Pezzoni, 2020; Zeppini and van Den Bergh, 2011). For example, preliminary evidence from Fusillo et al. (2020) showed that green technologies have a higher degree of diversity than their non-green counterparts, reinforcing the idea that green technologies should be considered complex due to the different bodies of technologies they combine. Further, Popp and Newell (2012) find that patents in sustainable energy domains are cited by a variety of other technological domains.

Consistently, innovation related to the SDGs is expected to be interdisciplinary and overcome the lack of holistic vision that often characterizes individual disciplines (Annan-Diab and Molinari, 2017). Although van der Waal et al. (2021) confirms that green technologies are relevant for many SDGs, no study attempts to quantitatively measure the interdisciplinarity of SDGs-related innovations, neither those more environmentally or more socially related.

To this end, in this study we follow the framework proposed by Rafols and Meyer (2010) who define diversity as the differences in the body of integrated knowledge that can be summarized by three attributes:

1. Variety: the number of different categories in which an element can be classified;

2. Balance: the evenness of the distribution of elements among categories;

3. Disparity: the degree of diversity between these categories.

The Rao-Stirling (RS) diversity index was originally proposed by Rao (1982) and then revised by Stirling (2007) to consider all the above three elements simultaneously, and it is often considered in the analysis of research interdisciplinarity, although its validity has been discussed in recent studies such as Leydesdorff et al. (2019).

In view of this, we put forward the first hypothesis:

Hypothesis 1 (H1): *SDGs related patents are more diverse than non-SDGs related ones across the different technological fields.*

3.2.3 The role of American universities for SDGs

Universities are fundamental actors in the innovation ecosystem, which are expected to have a "public mission" which consists of providing knowledge, critical thinking, and technological advances to tackle society's fundamental problems (Winickoff, 2013; Papadimitriou, 2020). In particular, since the Bayh-Dole Act, American universities have been encouraged to pursue the so-called "third mission", maximizing the societal benefit of technology they produce (Lemley, 2007). Therefore, the third mission of contemporary universities includes fostering the achievement of the SDGs (Lozano et al., 2013; Ceulemans et al., 2015; Blasco et al., 2020). Education, research and innovation are the three areas where universities play a vital role in putting society on the path of sustainable development (Körfgen et al., 2018; Leal Filho et al., 2019). While there are extensive studies in the literature about implementation of SDGs into university curricula (Albareda-Tiana et al., 2018; Álvarez et al., 2021; Sánchez-Carracedo et al., 2021; Thomas, 2016), academic research about the SDGs is expected to use a transformative approach which brings together different fields of study and uses interdisciplinarity as a "competitive advantage" to address complex societal challenges (Leal Filho et al., 2019). To this end, universities and research centers around the world have made significant progress towards establishing collaborative, interdisciplinary initiatives in sustainability science (Hernandez-Aguilera et al., 2021). This idea is consistent with the momentum that interdisciplinary research (IDR) is having in universities, producing wide-ranging scientific advances and leading to the establishment of interdisciplinary research centers (Biancani et al., 2018). Although interdisciplinarity research of SDGs is a trending topic (El-Jardali et al., 2018; Kestin et al., 2017), quantitative contributions on this topic are still scarce.

Further, the role of scientific research by universities to achieve the SDGs is not limited to a "knowledge phase", where research is conducted to answer the question "what is at present", but it includes a "technological phase" as well, where universities develop technological solutions to solve specific SDGs challenges (Kestin et al., 2017). Therefore, the research function of universities is closely linked to the production of innovation. Recently, universities have been shown to be particularly important in the development of environmental innovations because they accumulate a wide range of expertise and competencies that are distributed across different organisations (De Marchi, 2012; De Marchi and Grandinetti, 2013). Because of their particular educational endowments, inventors within universities are thought to have diverse knowledge bases and skills that enable them to successfully recombine bits of knowledge from different technological fields domains (Quatraro and Scandura, 2019).

Based on this conceptual background, we intend to analyze the production of innovation of American universities related to the SDGs, to check whether the interdisciplinary environment that contemporary universities are deemed to foster has an impact on the technological diversity of university produced innovation; thus, we put forward our second hypothesis:

Hypothesis 2 (H2): *University patents related to each of the SDGs are more diverse compared to other university patents.*

3.3 Research Design

3.3.1 Patent data collection

We use as a source of data for this research the information in patents granted at USPTO, considering them as a viable proxy to study the domain of technological domains in the knowledge economy (Jaffe and Trajtenberg, 2002). In order to use patents filed at USPTO, we relied on Patents View (Version 2021) where we collected all patents granted from 2006 to 2020. The Patents View platform is built on a regularly updated database that longitudinally links inventors, their organizations, locations, and overall patenting activity and reports the name(s) of first assignee(s), while the US Patent Assignment Database contains detailed information on patent assignments and other transactions recorded at the USPTO. The data collected from Patents View include general information about the patents (patent identification code, grant date), patent text (title, abstract and claims)[1], the IPC and WIPO classes associated with each patent. Information about patent quality is retrieved from OECD 2021 patent quality dataset.[2] Our analysis includes the specific subset of American universities' patents, defined as those having at least one university among the assignees in the patent history. After merging together the general information about patent text, the dataset has 85'169 patents including US universities among their applicants granted in the time range from 2006 to 2020.

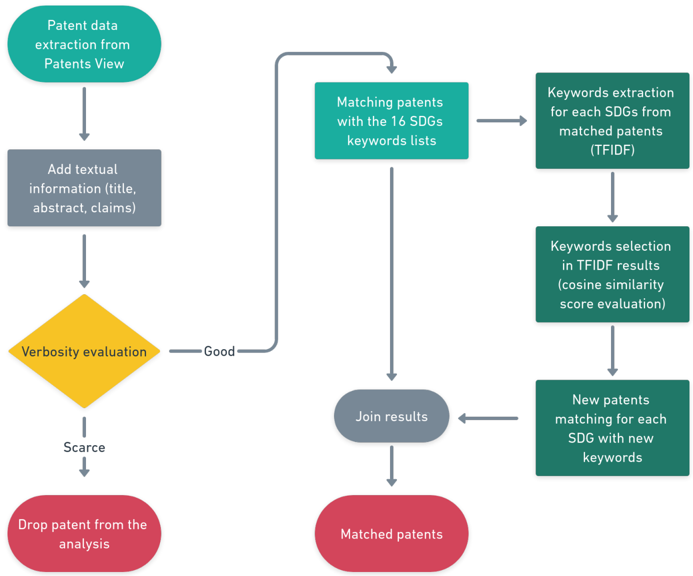
3.3.2 Tagging SDGs related patents

We approach the research questions exploring SDGs-relevant innovations using patent data and a textual analysis of their content. To identify patents related to interdisciplinary or integrated technologies or

[1] Two tables were used to collect the textual part of the patents. The first table is "patent.tsv" where it is possible to find information about abstract and title of patents; then, the claims information was retrieved joining each year's claim table from Patents View to the patents.tsv table. The data is free and accessible at https://patentsview.org/download/data-download-tables

[2] The OECD patent quality dataset (version 2021) is available upon request at https://www.oecd.org/sti/inno/intellectual-property-statistics-and-analysis.htm

Figure 3.1: Flowchart of SDGs related patent identification



Notes: The flowchart represents the research design of this work, emphasizing the four steps that compose the SDG tagging process: i) the first round of patent tagging using the original keywords lists, ii) the tf-idf keywords expansion, iii)the cosine-similarity based keywords selection iiii) the final round of patent tagging

emerging products, keyword search can be used as an effective method (Xie and Miyazaki, 2013). Patent documents are divided into several elements, including title, abstract, claims, and description, and since their purposes differ, their sentence structures and vocabulary also differ from each other. First, the title and abstract use distinctive and significantly differentiated words to properly express the relevant technologies, but are short and lack specific details about them. On the other hand, the claims are more complete and explicitly describe the related technical features to ensure complete legal protection, which is essential for patents (Noh et al., 2015). Previous work by van der Waal et al. (2021) and Hajikhani and Suominen (2021) also used patent text to map patents to the SDGs. In particular, the former use direct keyword matching from an initial keyword list and then label the green and "blue" (socially oriented) patents relevant to the SDGs; the latter propose a supervised machine learning approach instead: first, they create a dictionary of SDG-relevant keywords from UN SDGs documents, then they use it to identify relevant publications in the SCOPUS database. Second, they preprocess the text and convert it into a TF-IDF matrix, to create word embeddings to be used for classification. After validating the model, their classifier is trained on labeled publication data to predict the vector of probabilities that a patent belongs to each of the SDGs. However, they are not able to evaluate the quality of their patent results because they lack a *ground truth* on patent data (they only have pre-labeled data on publications and not on patents). In line with the previously mentioned literature, we perform a text analysis of the title, abstract and claims of patents, which are considered by the authors to be the most appropriate parts of the text for performing a quantitative analysis and a keyword search to avoid type I errors (missing patents that should be identified) and type II errors (retrieving irrelevant patents) (Xie and Miyazaki, 2013).

3.3.2.1 First round of matching and TF-IDF

As reported in Figure 3.1, the first step for tagging SDGs related patents consists of a direct matching between a *corpus* generated from joining each patent title, abstract and claims and 16 lists of keywords related to the SDGs (one for each SDG expect for SDG 17 which focuses on strengthening the partnerships to reach the other goals, thus is omitted from this analysis) developed by the University of Auckland SDGs keywords mapping research project [3]. The choice of this keyword list is due to its completeness, considering it combines Elsevier's keywords, a subset of Sustainable Development Solutions Network (SDSN) and UN keywords. Moreover, the list was generated using a text mining approach on academic publications, hence we selected it also for its consistency with the methodology proposed in this research.[4] Through this first round of matching, we are able to tag 426'863 patents related to at least one SDG as those that have at least one keyword from the corresponding list in their text.[5] However, considering that the keywords lists are based on academic publication text as well the peculiarities of patent texts and their specific legal jargon (especially in the claims) and the technical words that are not common in everyday language (Bonino et al., 2010; Tseng et al., 2007), we decided to perform a keyword extraction procedure to avoid Type I errors as much as possible. The criteria for selecting keywords from a document may also vary. For example, words that occur most frequently in certain documents may be considered critical, or words that fit well with the main topics of the document are often assumed to be important. In general, while words that occur frequently in patent documents are likely to be representative keywords, those that occur too frequently in such documents are also likely to be general words that occur in all documents rather than representative words that allow specific patents to be identified. Noh et al. (2015) sought to identify the most effective keyword strategy for text mining-based patent analysis by evaluating and verifying the most commonly used keyword selection and processing methods in existing studies. Their results highlight TF-IDF (Term Frequency-Inverse Document Frequency) as the best performer because it has the lowest entropy values. The term frequency (TF) is the number of times a term appears in a document and is calculated as follows:

$$tf_{ij} = \frac{n_{ij}}{|d_j|} \tag{3.1}$$

tf_{ij} stands for the number of occurrences of word i in document j and $|d_j|$ is the dimension, expressed by the number of words of j. Inverse Document Frequency (IDF) measures the rarity of a term in the whole

[3] More information on the project is available at https://www.sdgmapping.auckland.ac.nz/

[4] In particular, the list was made applying an n-gram model to mine the abstracts of academic publications, in order to identify relevant sequences of words. Afterwards, n-gram tokens were then scored by a range of factors, including counts and measures of frequency, and were then ranked by those scores. Keywords with a high rank were then evaluated in more detail and manually reviewed to confirm that they were relevant to the Goal in question.

[5] Considering the overlapping among SDGs (Nilsson et al., 2016), some keywords might be repeated in different lists.

corpus. It is calculated as follows:

$$idf_i = log_{10} \frac{|D|}{|d : i \epsilon d|} \tag{3.2}$$

where the denominator is the number of documents containing i. The concepts of term frequency and inverse document frequency are combined, to produce a composite weight for each term in each document, with the following formula.

$$tf - idf = tf_{ij} * idf_i \tag{3.3}$$

In this way, the TF-IDF method can retrieve important keywords that are closely related to a representative technology while avoiding general terms in the corpus (Usui et al., 2007). Thus, TF-IDF allows us to score each word in each patent document according to its weight both in the single patent text and in the whole collection of texts. For this analysis, we performed TF-IDF separately for each SDG, considering each SDG as a separate collection of texts.[6]

3.3.2.2 Keyword selection through cosine similarity and second round of matching

To evaluate the effectiveness of TF-IDF, we confronted the resulting extra keywords with the original lists. To make the comparison, we selected the 10 most relevant keywords for each patent according to their score and we compared them with the keywords in the original lists that had at least one match in any patent text. The results confirm the validity of TF-IDF: more than half of the keywords scored among the top 10 for each patent are also in the original SDGs lists, hence confirming the validity of this unsupervised method. Moreover, a further step of selection was needed to choose among the top 10 scored keywords per patent the most relevant ones to expand the original SDGs dictionaries. To do that, we took advantage of the vectorial representation of words elaborated by a pre-trained transformer based neural network (SentenceTransformers)[7], a kind of technique which has already demonstrated huge potentialities for patent analysis (Li et al., 2018; Chen et al., 2020; Roudsari et al., 2021). For each SDG, we compared the n-dimensional vector representing each keyword in the original list to the n-dimensional vector representing each word in the top 10 words per patent according to TF-IDF score. Among them, for each keyword in the original lists we selected the closest 3 in terms of cosine similarity.[8] Therefore, exploiting the vectorial representations of patents and their spatial and semantic closeness, as measured by cosine similarity, we are able to enlarge the original dictionaries with two different types of keywords:

---

[6]In this research, Python *scikit-learn* library is used to carry out the tf-idf because it considers also bi-grams and tri-grams that are more common than monograms in the original SDGs lists.

[7]SentenceTransformers (`https://www.sbert.net/`) is a modification of the pretrained BERT network that use siamese and triplet network structures to derive semantically meaningful sentence embeddings that can be compared using cosine-similarity. The main advantage of the model is that it is optimized to calculate cosine similarity, while maintaining the accuracy from BERT. Further information on the architecture is available in the work of Reimers and Gurevych (2019)

[8]Cosine similarity measures the cosine of the angle between two vectors, whose values range between -1 and 1. It is widely used as technique to assess semantic similarity between two documents, including patents' text. For instance see recent work by Rogers (2020) and Feng (2020)

1. Keywords which enrich concepts already present in the original list with semantically relevant keywords/synonyms (e.g., in the original list of SDG 9-Industry, Innovation and Infrastructure- there is the keyword "Sustainable Industrialization" whose closest keywords are "industrial waste" and "renewable", which are semantically close ideas; or, in SDG 3- Good Health and Wellbeing- the keyword "Sexual Health" has as closest keywords "sexual disorders" and "reproductive health").

2. Keywords which specify concepts and ideas already present in the original list with more technical wording (e.g., "Photochemistry" is a keyword present in the original list of SDG 7-Sustainable Development- and one of the closest identified keywords is "photocatalysis"; or in SDG 13-Climate Action- the bigram "greenhouse gas" has as closest keyword "CO2" which is the primary greenhouse gas emitted through human activities).[9]

Table 3.8 in the Appendix gives further detail about the results of the two rounds of matching. In particular, in the second round of matching, using the keywords derived from TF-IDF and selected through the cosine similarity method, we are able to add 1'784 keywords that produce a match in our patent record.[10] The total number of SDGs related patents identified is 693'571.[11]

3.3.3 SDGs related patents

Through the two rounds of matching, we are able to identify 693'571 SDGs related patents. Figure 3.2 depicts the trends in SDGs related patents granted from 2006 to 2020. The green line refers to the share of SDGs related university patents on the total of university patents for each year of the range, while the red line depicts the share of non-university SDGs related patents on the total of non-university patents. The share of university patents related to the SDGs is greater than the non-university counterpart, peaking at 33% in 2019, while patents related to the SDGs filed by other actors rise up to around 21%. Further, Table 3.1 presents the total number of SDGs related patents and the percentage of each SDG over the total of patents (Column 3) and the total of university patents (Column 6). Column 4 and Column 7 respectively report the composition of SDGs patents with respect to each SDG in the total number of patents and of university patents. The results presented in the Table 3.1 highlight the differences between production of SDGs related innovation in American universities and considering all US assignees. The biggest difference is represented by SDG 3 which is prominent in the total SDG landscape with a share of almost 15 %, but predominant in university innovation production, representing around 53 % of all

[9] For each SDG, the enlarged list of keywords is available upon request.

[10] Based on the frequencies of the matched terms we excluded 34 noisy keywords, derived from the TF-IDF. We decided, for each SDG, to exclude from the keywords those whose matching frequency was higher than the average frequency of all terms.

[11] As robustness check for this results, we compared the number of SDGs related patents to the number of green patents (as defined by two established international classifications, both based on the International Patent Classification (IPC): WIPO IPC Green Inventory (WIPO, 2012) and OECD ENV-TECH (Haščič and Migotto, 2015) and to the number of circular economy patents, as defined in Fusillo et al. (2021). In our data, almost 34% of green patents and 63% of circular economy patents are identified also as SDGs-related patents.

Table 3.1: Number of SDGs related patents and % by assignee type

SDG	all-assignees	% on tot.patents	%_all_SDG	univ-assignee	%_tot_univ	%_SDG univ
SDG1	66'998	1,8	9,7	1'093	1,3	4,3
SDG2	39'367	1,1	5,7	2'193	2,6	8,7
SDG3	101'129	2,8	14,6	13'486	15,8	53,4
SDG4	46'255	1,3	6,7	1'533	1,8	6,1
SDG5	135'821	3,7	19,6	3'712	4,4	14,7
SDG6	18'712	0,5	2,7	663	0,8	2,6
SDG7	85'761	2,4	12,4	2'801	3,3	11,1
SDG8	23'294	0,6	3,4	520	0,6	2,1
SDG9	26'271	0,7	3,8	552	0,6	2,2
SDG10	8'156	0,2	1,2	77	0,1	0,3
SDG11	235'204	6,5	33,9	2'387	2,8	9,5
SDG12	31'708	0,9	4,6	819	1	3,2
SDG13	12'426	0,3	1,8	255	0,3	1
SDG14	8'949	0,2	1,3	249	0,3	1
SDG15	10'211	0,3	1,5	358	0,4	1,4
SDG16	11'306	0,3	1,5	187	0,2	0,7

Notes: The table represents the number and percentages of patents for each SDG. In particular, the second column shows the total number of patents for each SDG considering all the patents in our sample. The third column shows the percentages of each SDG patents on the total number of patents. The fourth column instead shows the percentage of each SDG on the total of all SDGs patents. The last three columns respectively represents the total number of each SDG related patent in university patents, the percentage on the total number of university patents and eventually the percentage of each SDG on the total number of university SDG patents. The sum of the percentages are greater than the unity because each patent can be assigned to more than one SDG at the same time.

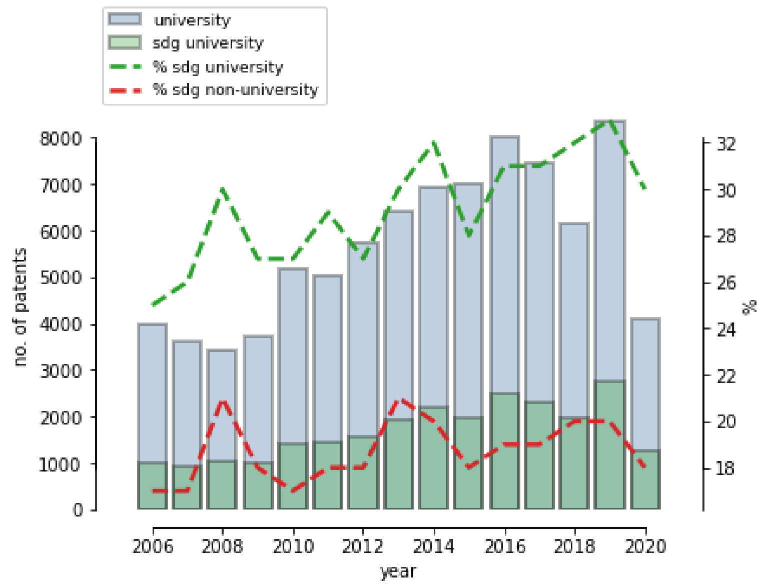
university patents related to the SDGs. Universities also seem to focus more on SDG 2 (almost 9 % vs 6 % in all patents). This figure is to some extent consistent with the results of van der Waal et al. (2021) who find, using EPO patent data, that most important contribution of multinational corporations is related to SDG 3. Figure 3.3 represents the percentage of patents, for each SDG, belonging to the top 10 most common IPC codes. We observe that, on the one hand, out of the 10 codes, only 4 are identify green patents, namely H01L, H01M, C12N and G06Q according to the classification of the WIPO *Green Inventory*. It is to be noted that green IPC codes stand out more in SDGs related to environmental innovation. On the other hand, most of the SDGs seem to be associated with the G06F class which is related to computer systems based on specific computational models and with the H04L class which covers the transmission of digital information. Both of these classes belong to the high-tech IPC classification[12]. Another relevant technological class for some of the SDGs, especially for SDG 3, is A61K, which refers to medical preparations and pharmaceutical products. These results are especially interesting considering that currently the IPC classes related to sustainable development are limited to green technologies and do not cover "improved conditions" and social sustainability, which is currently possible to identify through semantic search.

3.3.3.1 Universities' SDGs related patents

This research has a special focus on universities' SDGs innovation. To this end, we identify 25'247 SDGs-related university owned patents. Figure 3.4 is a network where each of the 25'247 university

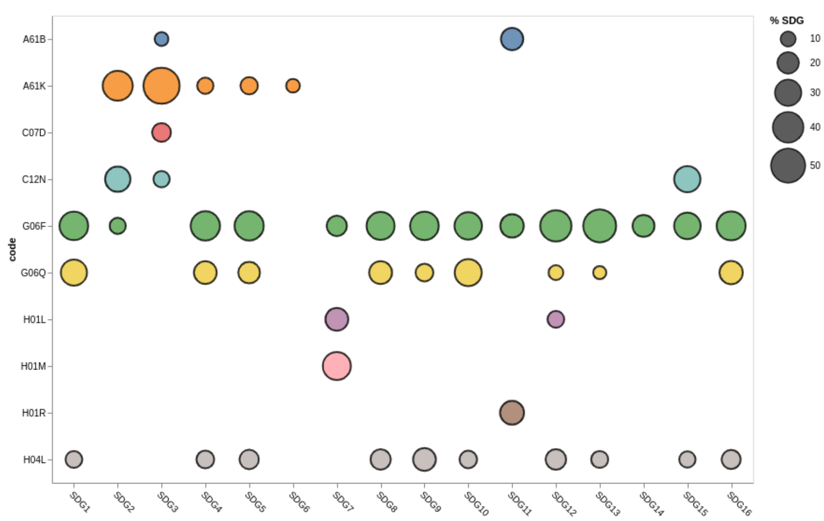
[12]The definition of high-technology patents proposed by Eurostat uses specific subclasses of the International Patent Classification (IPC) as defined in the trilateral statistical report of the EPO, JPO and USPTO. The list is accessible at https://ec.europa.eu/eurostat/cache/metadata/Annexes/pat_esms_an2.pdf

Figure 3.2: Trends in SDGs related patents



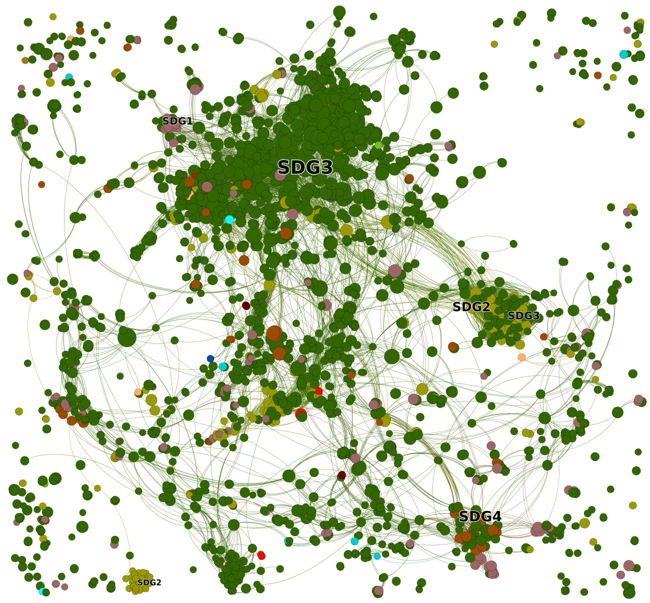
Notes: This distribution is related to all university owned patents in the time range 2006-2020.

Figure 3.3: Top 10 most comment 4-digit IPC codes for SDGs related patents



Notes: The figure represents the percentage of patents, for each SDG, belonging to the top 10 most frequent 4-digit IPC codes in the total distribution. The bubbles are proportional to the percentage and each color is linked to a specific IPC code.

Figure 3.4: SDGs cosine similarity network (university patents only)



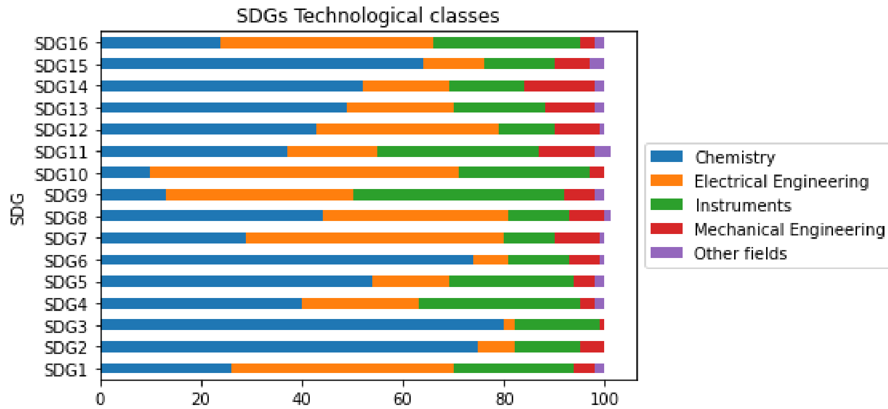
Notes: The figure represents a section of a network where each patent of the 25'247 identified SDGs university patents is a node and the edges' length is proportional to the cosine distance between each couple of patents. For the purpose of clarity, we plot only the edges where the cosine similarity is above a threshold of 0.6.

Table 3.2: Top 15 universities for share of SDGs related patents

University	**No. patents SDGs**	**No. patents**	**% SDG**
Johns Hopkins University	620	1'691	37
University of South Florida	427	1'145	37
New York University	339	925	37
Duke University	314	897	35
University of North Texas	562	1'661	34
University of Florida	323	1'038	31
Ohio university	842	2'787	30
Cornell University	359	1'185	30
University of Illinois	290	967	30
University of California	2'136	7'500	28
Michigan State University	345	1'269	27
Stanford University	677	2'597	26
University of Wisconsin Madison	375	1'530	25
Massachusetts Institute of Technology	733	3512	21
California Institute of Technology	337	1'945	17

Notes: The ranking is based on the number of USPTO patents held by American universities in the 2006-2020 time span.

Figure 3.5: SDGs technological class composition



Notes: The graph represents the technological composition in percentage of each SDG (SDG 1 to SDG 16) according to the five WIPO technological classes (Chemistry, Electrical Engineering, Instruments, Mechanical Engineering and Other fields).

SDG-related patents is a node and the edges are weighted according to the cosine similarity among the vectorial representation of each patent. From the graph it is clear that university SDGs innovation production mostly revolves around SDG 3 and SDG 2 and, interestingly, the patents related to both are semantically closer one another. This might be a hint about the diversity of these patents, which will be verified through the econometric estimations in Section 3.3.6. In addition, in the Appendix, it is possible to find the most frequent keywords associated with each SDG. Table 3.2 reports the top 15 universities in terms of SDGs related patents in the time span 2006-2020. In absolute numbers, the best performer is University of California with 2'136 SDGs related patents. However, looking for universities with the highest share of SDGs related patents the best performer are the Johns Hopkins University, the University of South Florida and the New York University with a share of around 37% SDGs related patents. Figure 3.5 shows the percentage of the 5 WIPO classes composing each of the 16 SDGs. Chemistry seems to be the prevailing class in almost all the SDGs, especially for SDG 2 (Zero Hunger), SDG 3 (Good Health and Wellbeing) and SDG 6 (Clean water and Sanitation) where the Chemistry share is around 75%. This seems reasonable, considering health and sanitation are both fields where chemical technologies are fundamental. Furthermore, SDG 7 (Affordable and Clean Energy) and SDG 9 (Industry, Innovation and Infrastructure) have the greatest share of Electrical Engineering patents, around 50%. This is consistent with the fact that, being focused on innovation, SDG 9 includes many ICT related technologies. At the same time, ICT technologies seem to be relevant also for more socially oriented goals, such as SDG 16 (Peace, Justice and Strong Institutions) and SDG 8 (Decent work and Economic Growth).

3.3.4 Technological diversity of patents through Rao-Stirling index

As mentioned before, interdisciplinarity can be decomposed in three main components: variety, balance and disparity (Leydesdorff, 2018; Wang et al., 2015), which are accounted for in the Rao-Stirling index.

The Rao-Stirling index has been widely used to measure diversity and in general interdisciplinarity (Porter et al., 2007; Porter and Rafols, 2009). The indicator is defined as follows:

$$\Delta = \sum_{ij} p_i p_j d_{ij} \tag{3.4}$$

where d_{ij} is a measure of cognitive distance between classes i and j and p_i and p_j are the proportions of elements assigned respectively to class i and j. Considering patents as our work unit, the classes are, in this case, 4-digits-IPC codes to which patents are co-assigned. Thus, the diversity of each patent is calculated as the proportion of IPC4 codes weighted by their cognitive distance.

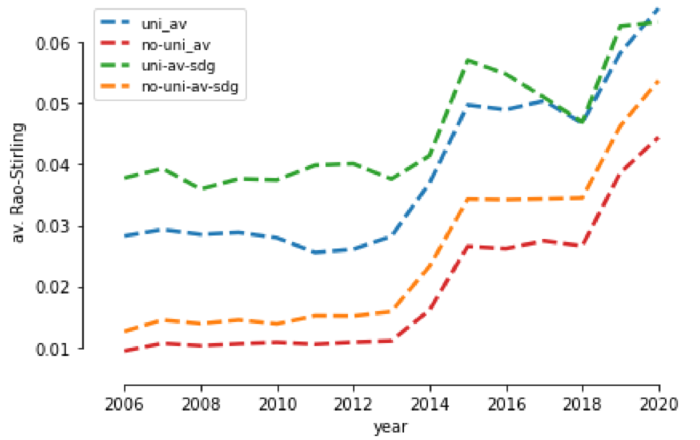
The first step is calculating the cognitive distance. This research uses co-classification, the frequency with which two classes are assigned to a patent, as a measure of cognitive distance (defined also as 'proximity'). Considering a patent usually belongs to multiple classes, scholars have often used co-classification of patents to develop indicators of distance among technological fields (Engelsman and van Raan, 1994). The underlying assumption is that if the frequency with which two classes are jointly assigned is high, these two classes are proximate (Yan and Luo, 2017). Thus, we use co-occurrences between IPC4 codes as a measure of cognitive distance, creating a symmetric co-occurrence matrix C in which each term C_{ij} represents the number of patents linked to IPC class i and j.

However, the recent work by Alstott et al. (2017) highlighted that all the measures of technology proximity might be affected by factors other than the technologies themselves. In the case of measures of co-occurrence, the authors claim that the simple number of occurrences can be considered as an impinging factor. In particular, the probability that two classes co-occur in the same patent depends on the number of classes that are associated with a patent and the number of patents that are associated with a certain technological class. Bottazzi and Pirino (2010) propose to overcome this issue comparing the observed co-occurrence against the null hypothesis in which the co-occurrences of classes are randomly distributed, preserving, at the same time, both the number of occurrences of a class and the number of classes that are associated with the selected patent. Alstott et al. (2017) propose to verify this null hypothesis creating 1000 randomized control matrices in which the number of occurrences of each class and the number of patents per class are preserved. However, the two constraints proposed by Bottazzi and Pirino (2010) are consistent with the moments (such as mean and standard deviation) of a hypergeometric distribution as explained in the seminal work of Teece et al. (1994). Hence, under the assumption of joint random occurrences and hypergeometric distribution, the mean and the standard deviation are calculated as following:

$$\mu = \frac{C_i C_j}{N} \tag{3.5}$$

$$\sigma_{ij}^2 = \mu_{ij}\left(\frac{N - C_i}{N}\right)\left(\frac{N - C_j}{N - 1}\right) \tag{3.6}$$

Figure 3.6: Rao-Stirling index



Notes: The graph represents the yearly average of Rao-Stirling index for patents granted between 2006 and 2020.

Where N is the number of patents, C_i and C_j the numbers of patents respectively linked to class i and class j. The relation between class i and j can be express through a z-score, where C_{ij} is the empirical co-occurrence value between class i and j:

$$r_{ij} = \frac{C_{ij} - \mu_{ij}}{\sigma_{ij}} \tag{3.7}$$

Calculating term r for every couple of IPC classes we obtain a matrix R_{ij} and using a cosine normalization, we obtain a S_{ij} matrix whose values range from 0 to 1. At this point, it is possible to calculate the cognitive distance for each patent, through the following:

$$d_{ij} = 1 - s_{ij} \tag{3.8}$$

whose result is used to weight the proportion of technologies in Rao-Stirling index.

3.3.5 Technological diversity assessment

Figure 3.6 represents the average of Rao-Stirling index, calculated as previously mentioned and plotted for the time range 2006-2020. The green line represents the trend of university patents related to SDGs, while the blue line represents university patents not related to SDGs. University patents (blue and green lines) are, on average, more diverse compared to non university patents (red and orange lines). This might be explained considering that diverse innovation, such as the one required by SDGs, might be better tackled by collective efforts of universities and research institutions where teams of inventors collaborate to generate innovation rather than individual inventors (Quatraro and Scandura, 2019; Orsatti, Quatraro and Pezzoni, 2020). Furthermore, patents related to SDGs seem to be slightly more diverse compared to non-SDG related patents for both university and non-university patents. In particular, university patents perform the best in terms of diversity from 2015 (the year of the adoption of the Agenda for Sustainable Development and of the SDGs). These findings will be further explored through the econometric models presented in the following section.

3.3.6 Empirical Strategy

3.3.6.1 Diversity of SDGs vs non-SDGs related patents across different technological fields

In order to empirically test Hypothesis (1) we estimate the following model:

$$\begin{aligned} \Delta_i = \alpha + \beta_1 SDG_i + \beta_2 inventors_i + \beta_3 familySize_i \\ +\beta_4 backCits_i + \beta_5 claims_i + \beta_6 univ_i + IPC.3digit_i + t_i + w_i + \epsilon_i \end{aligned} \tag{3.9}$$

The dependent variable is Δ_i, the Rao-Stirling diversity index and our focal explanatory variable is SDG_i, taking value 1 if the patent is related to SDGs and 0 otherwise. The model includes controls for the number of claims ($claims_i$) and the number of backwards citations ($backCits_i$), as well as for the number of distinct inventors ($inventors_i$), family size ($familySize_i$), which can be linked to the degree of diversity, and for the fact of being owned by a university ($univ_i$). To account for time varying effects we include a set of 15 dummies for the 15 years span considered (t_i) and a set of state controls to account from geographic heterogeneity within the US (w_i); to account for technological heterogeneity instead, we add narrow technological controls ($IPC.3digit_i$) which is a set of IPC 3-digit dummy variables that capture the specific features of each technological domain. We include the latter control following the suggestion of Barbieri et al. (2020) to increase the robustness of the analysis by eliminating the risk of the coefficient of SDG variable being driven by effects that are related to the different technological fields. The estimation could be biased by ignoring the peculiarities of technological areas, such as the availability of consolidated previous work and the inclination to rely on a broader knowledge base. However, it should be noted that adding these dummies, limits the analysis to those IPC 3-digit codes that include at least one SDG and one non-SDG patent.

Furthermore, all reported standard errors are heteroskedastic robust. Considering our dependent variables ranges between 0 and 1 and there is no consensus about the econometric specification for this kind of dependent variable, the best choice seems to carry out the estimation through OLS regression (Fusillo et al., 2020). Furthermore, we take into account that the sample of considered patents might be highly heterogeneous in terms of patent characteristics and to partially tackle this issue, the analysis is conducted considering separately the 5 WIPO technological macro field (Chemistry, Mechanical Engineering, Instruments, Electrical Engineering and Other).

3.3.6.2 Diversity of university patents related to the SDGs

After testing for the technological diversity of SDGs related patents, we specifically focus on American universities patent portfolios, to check whether university patents related to each SDG have a diversity premium. Thus, in order to empirically test Hypothesis (2) we estimate the following model using university owned patents only:

$$\begin{aligned}\Delta_i = \alpha + \beta_1 SDG1_i + \beta_2 SDG2_i + \beta_3 SDG3_i + \beta_4 SDG4_i + \beta_5 SDG5_i + \beta_6 SDG6_i + \beta_7 SDG7_i \\ +\beta_8 SDG8_i + \beta_9 SDG9_i + \beta_{10} SDG10_i + \beta_{11} SDG11_i + \beta_{12} SDG12_i + \beta_{13} SDG13_i + \beta_{14} SDG14_i \\ +\beta_{15} SDG15_i + \beta_{16} SDG16_i + \beta_{17} inventors_i + \beta_{18} familySize_i \\ +\beta_{19} backCits_i + \beta_{20} claims_i + \beta_{21} renewal_i + t_i + w_i + \epsilon_i\end{aligned} \tag{3.10}$$

The dependent variable is Δ_i, the Rao-Stirling diversity index and each of the focal explanatory variables ($SDG1_i$ to $SDG16_i$) is a dummy taking the value 1 if the patent is related to corresponding SDG. The controls are the same as presented in Section 3.3.6.1.

3.4 Results

3.4.1 Technological diversity of SDGs-related patents

Section 3.2.2 and Section 3.2.3 put respectively forward the two hypotheses of the present research. The first is that SDGs related patents are more diverse than non-SDGs related ones across the different technological fields. The second hypothesis is that university patents related to the SDGs are more diverse compared to other university patents. Table 3.3 presents the descriptive statistics. Further, Table 3.9 in Appendix presents the pairwise correlations.

Table 3.4 presents the results related to the first hypothesis. Columns 1, 2, 3, 4 and 5 report the estimates for the 5 WIPO technological sectors. The SDG coefficient is positive and signficant in all the cases expect for the Chemistry field where it is significant and negative. Thus, this first set of results mostly confirm our first hypothesis, showing that SDGs related technologies are more diverse as compared to non-SDGs related inventions.

Table 3.5 presents the results related to the second hypothesis, shedding light on the interplay between university patents, technological diversity and SDGs. Considering SDGs cover wide and distinct areas of knowledge and technologies, it is worth observing them separately. The results show that, when looking at university patents only, the diversity effect is heterogeneous across different SDGs. On the one hand, SDG 2, SDG 3, SDG 4 and SDG 15 have a positive and significant coefficient; on the other hand, SDG 1, SDG 6, SDG 7, SDG 9, SDG 10 and SDG 14 have a significant and negative coefficient, while SDG 5, SDG 8, SDG 11, SDG 12, SDG 13 and SDG 16 have a non significant coefficient.

Table 3.3: Descriptive statistics (2006-2020)

	Obs	**Mean**	**SD**	**Min**	**Max**
Rao Stirling	3’640’513	.022	0.059	0	.386
SDG	3’640’513	.191	0.393	0	1
SDG 1	3’640’513	.018	0.134	0	1
SDG 2	3’640’513	.011	0.103	0	1
SDG 3	3’640’513	.028	0.164	0	1
SDG 4	3’640’513	.013	0.112	0	1
SDG 5	3’640’513	.037	0.190	0	1
SDG 6	3’640’513	.005	0.072	0	1
SDG 7	3’640’513	.024	0.152	0	1
SDG 8	3’640’513	.006	0.080	0	1
SDG 9	3’640’513	.007	0.085	0	1
SDG 10	3’640’513	.002	0.047	0	1
SDG 11	3’640’513	.065	0.246	0	1
SDG 12	3’640’513	.009	0.093	0	1
SDG 13	3’640’513	.003	0.058	0	1
SDG 14	3’640’513	.002	0.050	0	1
SDG 15	3’640’513	.003	0.053	0	1
SDG 16	3’640’513	.003	0.056	0	1
Family Size	3’303’762	3.747	3.788	1	57
Inventors	3’640’513	2.759	1.944	1	133
Backward cits.	3’303’762	27.603	75.379	0	858
Claims	3’303’728	16.708	10.580	1	803

Overall, these results might be aligned with the literature which defines technologies aiming to tackle the multifaceted issue of our century as more diversified (Fusillo et al., 2020) and the idea that the promotion of interdisciplinarity within higher education institutions has become widespread over the last few decades (Tarrant and Thiele, 2017). However, these results only partially confirm our second hypothesis, considering there is a diversity premium only for university patents linked to SDG 2, SDG 3, SDG 4 and SDG 15.

3.4.2 Robustness checks

To check the robustness of our empirical study, we present some additional estimates in this subsection. Nevertheless, we assume that the direction of the previously identified relationships should be robust to the choice of diversity measure, as long as these such measures are intended to capture the diversity construct. Therefore, we further check the robustness of our analysis by using alternative measures of technological diversity. To this end, we choose the index of technological diversity developed by Blau (1977) and recently adopted by Zhu et al. (2022) to measure the degree of knowledge recombination in a patent application. The indicator is defined as follows:

$$\text{Technological Diversity} = 1 - \sum_i \left(\frac{\text{no. of IPC}_i \text{ codes}}{\text{tot IPC codes}} \right)^2 \tag{3.11}$$

Table 3.4: OLS regression results of SDG on Rao-Stirling with time state and IPC controls

	(1) Chemistry	(2) Elec.Eng	(3) Instr.	(4) Mech.Eng	(5) Other
SDG	-0.0024***	0.0003***	0.0040***	0.0027***	0.0020***
	(0.0003)	(0.0001)	(0.0003)	(0.0003)	(0.0004)
family_size	-0.0004***	0.0002***	0.0004***	0.0005***	0.0001*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
bwd_cits	0.0000***	0.0000***	-0.0000***	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
claims	0.0001***	0.0000***	0.0000***	-0.0000	0.0000*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
no_inv	-0.0000	-0.0001***	0.0002***	0.0004***	-0.0001
	(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0001)
university	0.0010**	0.0018***	0.0035***	0.0048***	0.0067***
	(0.0004)	(0.0003)	(0.0005)	(0.0012)	(0.0022)
_cons	0.1251	0.0386***	0.0460***	0.0122***	0.0203***
	(0.0927)	(0.0041)	(0.0030)	(0.0019)	(0.0038)
Year Dummies	YES	YES	YES	YES	YES
State Dummies	YES	YES	YES	YES	YES
IPC.3digit	YES	YES	YES	YES	YES
Observations	264726	872576	303690	248920	104621
R^2	0.182	0.156	0.133	0.121	0.140

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: The dependent variable is the Rao-Stirling index as defined by Rao (1982). Unit of observation: patent. Grant years: 2006-2020. Heteroskedastic-Robust standard errors in parentheses; ***p<0.01,**p<0.05, *p<0.1

where IPC_i represents each unique 4-digit IPC code assigned to a patent. Further, considering our dependant variable is a continuous variable between 0 and 1, we use a Fractional response model which can be used to model a variable that takes values within a bounded range; the dependent variable may be any continuous variable bounded between 0 and 1, so that: $0 \leq y_i \leq 1$ (Papke and Wooldridge, 1996). The results of the robustness checks are presented in Table 3.6 and in Table 3.7.[13] In the former, we observe that the previous results hold even if we use as dependent variable the technological diversity ($tech\_div$) as defined by Blau (1977). Further, the latter confirms our results for SDG 1, SDG 2, SDG 3, SDG 8, SDG 12, SDG 13 and SDG 16; SDG 5, SDG 11 and SDG 15 become significant keeping the same sign while SDG 6, SDG 7 and SDG 14 take the opposite sign; finally, SDG 4, SDG 9 and SDG 10 lose their significant effect.

Further, considering literature suggests that green technologies are more diverse (Quatraro and Scandura, 2019; Fusillo et al., 2020), we check whether the premium diversity we observe is only due to environmental related SDGs technologies. To this end, in the Appendix we present the results of econometric specification where the independent variable SDG_i is split in the three components: environmental, social and development related. As shown in Table 3.10, the results confirm the diversity premium for all the three subgroups.

[13]However, the coefficients of Fractional response model cannot be easily interpreted by themselves and to ease interpretation, elasticities should be calculated.

Table 3.5: OLS regression results of SDG on Rao-Stirling with time and state controls (university patents only)

	(1) Rao-stirling	(2) Rai-stirling	(3) Rao-stirling
SDG 1	-0.0095***	-0.0107***	-0.0103***
	(0.0021)	(0.0021)	(0.0021)
SDG 2	0.0055***	0.0063***	0.0066***
	(0.0017)	(0.0018)	(0.0018)
SDG 3	0.0140***	0.0134***	0.0133***
	(0.0007)	(0.0008)	(0.0008)
SDG 4	0.0068***	0.0086***	0.0088***
	(0.0021)	(0.0022)	(0.0022)
SDG 5	-0.0004	-0.0004	-0.0003
	(0.0013)	(0.0013)	(0.0013)
SDG 6	-0.0145***	-0.0136***	-0.0136***
	(0.0026)	(0.0027)	(0.0027)
SDG 7	-0.0116***	-0.0128***	-0.0126***
	(0.0013)	(0.0013)	(0.0013)
SDG 8	0.0003	-0.0002	0.0000
	(0.0032)	(0.0033)	(0.0033)
SDG 9	-0.0157***	-0.0162***	-0.0155***
	(0.0025)	(0.0025)	(0.0025)
SDG 10	-0.0183**	-0.0153*	-0.0148*
	(0.0072)	(0.0080)	(0.0080)
SDG 11	-0.0008	-0.0011	-0.0010
	(0.0016)	(0.0017)	(0.0017)
SDG 12	0.0006	0.0004	0.0005
	(0.0026)	(0.0026)	(0.0026)
SDG 13	0.0039	0.0011	0.0012
	(0.0049)	(0.0049)	(0.0049)
SDG 14	-0.0076*	-0.0091**	-0.0090**
	(0.0045)	(0.0046)	(0.0046)
SDG 15	0.0265***	0.0301***	0.0305***
	(0.0050)	(0.0052)	(0.0053)
SDG 16	-0.0025	-0.0008	-0.0006
	(0.0054)	(0.0056)	(0.0056)
family_size		0.0008***	0.0007***
		(0.0001)	(0.0001)
no_inv		0.0004**	0.0004**
		(0.0002)	(0.0002)
bwd_cits			0.0000***
			(0.0000)
claims			-0.0002***
			(0.0000)
_cons	0.0259***	0.0227***	0.0264***
	(0.0028)	(0.0029)	(0.0030)
Year Dummies	YES	YES	YES
State Dummies	YES	YES	YES
Observations	84488	76957	76957
R^2	0.034	0.032	0.033

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: The dependent variable is the Rao-Stirling index as defined by Rao (1982). Unit of observation: patent. Grant years: 2006-2020. Heteroskedastic-Robust standard errors in parentheses; ***p<0.01,**p<0.05, *p<0.1

Table 3.6: Fractional regression results of SDG on Technological Diversity with time and state and IPC controls

	(1) Chemistry	(2) Elec.Eng	(3) Instr.	(4) Mech.Eng	(5) Other
SDG	0.0075	0.0187***	0.0194***	0.0693***	0.0775***
	(0.0055)	(0.0040)	(0.0069)	(0.0072)	(0.0132)
family_size	0.0033***	0.0252***	0.0122***	0.0137***	0.0061***
	(0.0004)	(0.0005)	(0.0008)	(0.0010)	(0.0017)
bwd_cits	-0.0003***	0.0002***	-0.0003***	0.0002***	0.0001
	(0.0000)	(0.0000)	(0.0000)	(0.0001)	(0.0001)
claims	0.0017***	0.0007***	0.0019***	0.0011***	0.0033***
	(0.0002)	(0.0002)	(0.0003)	(0.0004)	(0.0006)
no_inv	0.0317***	-0.0013	0.0111***	0.0190***	0.0210***
	(0.0010)	(0.0008)	(0.0014)	(0.0017)	(0.0031)
university	0.0706***	0.0993***	0.1712***	0.2188***	0.2516***
	(0.0072)	(0.0112)	(0.0114)	(0.0232)	(0.0507)
_cons	-1.2462	-0.8497***	-0.7092***	-1.8281***	-1.6391***
	(0.9112)	(0.0664)	(0.0549)	(0.0411)	(0.1037)
Year Dummies	YES	YES	YES	YES	YES
State Dummies	YES	YES	YES	YES	YES
IPC3.digit	YES	YES	YES	YES	YES
Observations	264726	872576	303690	248920	104621
pseudo R^2	0.062	0.121	0.120	0.133	0.153

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: The dependent variable is the Technological Diversity as defined by Blau (1977). Unit of observation: patent. Grant years: 2006-2020. Robust standard errors in parentheses; ***p<0.01,**p<0.05, *p<0.1

3.5 Discussion and Conclusion

Innovation for sustainable development plays a fundamental role in achieving the SDGs by fostering the creation of inventions that can help address the complex challenges of this century, such as environmental and social crises, and improve people's lives through advances in relevant sectors such as energy, medicine, and transportation (Blohmke, 2014; Bannerman, 2020; Rimmer, 2018).

However, the potential contribution of intellectual property to advancing the SDGs does not appear to have been explored in depth by scholars and practitioners. In particular, the World Intellectual Property Organization (WIPO) has only acknowledged the link between SDG 9 and IP (WIPO, 2018), while making a stronger contribution to the mapping of green technologies through the IPC codes based *Green Inventory*. Nevertheless, as we show in this research, green technologies are only a partial response to the challenges of SDGs (van der Waal et al., 2021).

In this work, starting from an initial NLP-derived keywords list, we used a robust, unsupervised methodology, to create a patent-related enriched dictionary that references 16 of the SDGs, allowing us to quantify interest in patents related to the SDGs and identify the most represented technology areas. This is a first contribution of this research, as no such patent-related dictionary has been proposed so

Table 3.7: Fractional regression results of SDG on Technological Diversity with time and state controls (university patents only)

	(1) Tech_div	(2) Tech_div	(3) Tech_div
SDG 1	-0.1315***	-0.1432***	-0.1411***
	(0.0411)	(0.0439)	(0.0439)
SDG 2	0.1349***	0.1304***	0.1311***
	(0.0263)	(0.0286)	(0.0286)
SDG 3	0.1016***	0.0973***	0.0910***
	(0.0121)	(0.0131)	(0.0132)
SDG 4	-0.0260	-0.0094	-0.0082
	(0.0342)	(0.0364)	(0.0365)
SDG 5	-0.1773***	-0.1848***	-0.1840***
	(0.0228)	(0.0246)	(0.0246)
SDG 6	0.1417***	0.1647***	0.1626***
	(0.0506)	(0.0556)	(0.0556)
SDG 7	0.1057***	0.1159***	0.1162***
	(0.0238)	(0.0255)	(0.0255)
SDG 8	0.0007	-0.0067	-0.0072
	(0.0548)	(0.0574)	(0.0574)
SDG 9	-0.0213	-0.0309	-0.0305
	(0.0536)	(0.0581)	(0.0582)
SDG 10	-0.0262	-0.0790	-0.0724
	(0.1391)	(0.1620)	(0.1622)
SDG 11	-0.1053***	-0.1033***	-0.0988***
	(0.0277)	(0.0300)	(0.0300)
SDG 12	0.0443	0.0313	0.0319
	(0.0449)	(0.0477)	(0.0478)
SDG 13	0.0414	0.0498	0.0480
	(0.0725)	(0.0773)	(0.0773)
SDG 14	0.2247***	0.2104**	0.2114**
	(0.0804)	(0.0861)	(0.0859)
SDG 15	0.1035	0.1276*	0.1252*
	(0.0704)	(0.0758)	(0.0759)
SDG 16	-0.0090	-0.0249	-0.0224
	(0.0919)	(0.1008)	(0.1007)
family_size		0.0126***	0.0142***
		(0.0011)	(0.0011)
no_inv		0.0194***	0.0202***
		(0.0026)	(0.0026)
bwd_cits			-0.0006***
			(0.0001)
claims			-0.0010**
			(0.0004)
_cons	-1.3492***	-1.4682***	-1.4494***
	(0.0509)	(0.0535)	(0.0541)
Year dummies	YES	YES	YES
State dummies	YES	YES	YES
Observations	84488	76957	76957
pseudo R^2	0.049	0.051	0.051

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: The dependent variable is the Technological Diversity as defined by Blau (1977). Unit of observation: patent. Grant years: 2006-2020. Robust standard errors in parentheses; ***p<0.01,**p<0.05, *p<0.1

far either by the literature or by commercial tools. These dictionaries might serve as a starting point for extracting other sustainability-related information from patent texts, thus improving the use of this type of data, which is not normally intended to facilitate communication about sustainability among stakeholders (Abrahamson and Baumard, 2008).

In addition, semantic-based patent analysis allows us to evaluate and compare which technical areas (i.e., IPC codes) contribute more to achieving the SDGs. This information could not be retrieved if considering only the technological classes of patents, since specific classes related to the Sustainable Development Goals have not yet been proposed (van der Waal et al., 2021). Our results shed light on the fact that green technologies only partially contribute to the achievement of the SDGs, while the predominant role is played by high technologies, whose contribution is in fact hardly recognized in the literature (Kostoska and Kocarev, 2019; Vinuesa et al., 2020). Therefore, this finding calls for a more careful consideration and understanding of the role of digital technologies in achieving the Sustainable Development Goals.

Second, this research also provides information on the role of US universities in producing innovation related to the SDGs, a role that is consistent with their mission to maximize societal benefits from the innovations they produce (Papadimitriou, 2020). In this context, we note that the filing of SDG-related patents by universities is increasing at a faster rate compared to other actors. This could be interpreted as universities becoming more aware that part of their public mission is to support the realization of the SDGs (Owens, 2017; Nilsson et al., 2016). However, university-generated innovations related to the SDGs are not evenly distributed among them: SDG 3 related patents account for more than 50% of the total number of SDGs-related patents from universities. This finding is consistent with van der Waal et al. (2021) results and is explicable considering the increasing importance of disease control in our society, especially after COVID19. At the same time, it could give rise to further debate and analysis, because patenting in the pharmaceutical field is not without negative consequences and criticism, especially when the university is the owner of the patent (Sampat, 2020, 2021). With this in mind, universities that are heavily involved in patenting for SDG 3 might consider adopting specific licensing strategies to maximize the associated social benefits, such as requiring licensing companies to distribute the product in developing countries beforehand or setting a fair price for the product at the time of technology transfer (Nelsen, 2002).

Third, this research has shown that, consistently with the literature on green innovation (Fusillo et al., 2020; Quatraro and Scandura, 2019), most techonologies related to the SDGs are more technologically diverse, confirming the first research hypothesis. This finding calls for a consistent policy intervention to better stimulate technological diversity to increase technological progress related to the SDGs and reduce environmental and social pressures.

This is especially true for universities, which are uniquely positioned to lead cross-sectoral implementation of the SDGs and provide an invaluable source of expertise in research and education on all areas of the SDGs (Owens, 2017; Nilsson et al., 2016). Moreover, it is increasingly important for universities to demonstrate not only to their financiers but to all stakeholders their ability to generate positive impacts on the territory through their coupling strategies. In this sense, it is critical to understand how the university generates societal benefits and how research activities impact societal benefits. Thus, the results of this study are intended to provide additional data that can help inform how university research projects generate societal impact. Indeed, interpolation between universities and SDGs does not always lead to a diversity premium for university patents. From this study, universities patents show a diversity premium only for SDG 2, SDG 3, which is the most prevalent in university SDG innovation production, and SDG 15. Thus, most university patents related to the SDGs do not show a diversity premium, even those related to greener SDGs, in contrast with what one would expect from the literature (Barbieri et al., 2020; Quatraro and Scandura, 2019). To improve this situation, we believe that specific investments in research and development that foster interactions between different disciplines and abet the creation of new knowledge-based networks are necessary to increase the diversity of SDG-related patents. In addition, to encourage the recombination of knowledge, it should be easier to obtain funding for interdisciplinary research than for mainstream activities (Rylance, 2015).

However, incentivizing interdisciplinarity does not come without any *caveat*. For instance, recent work by Zhu et al. (2022), showed that patent filings related to high diverse innovation run the risk of having a delayed patent grant, due to the increased complexity and ambiguity for patent examiners. For these reasons, the process to create highly diverse technologies should be carefully monitored, because delays in this kind of patents would entail delay in solving the *grand challenges* they are supposed to tackle.

This study is not free from limitations. First, patent data may not be able to fully capture technological innovation in the context of the SDGs, as it only partially represents a broader range of knowledge and technologies needed to promote sustainable development. Second, this research is based on initial keyword lists that influence the overall results and risk being incomplete or unbalanced across the SDGs, which may lead to an overemphasis or overweighting of some SDGs over others. Further, through this methodology is not possible to identify all SDGs-related patents whose text do not hint at SDGs keywords; thus, we fail to identify this kind of patents (false negative cases). This limitation is especially relevant for the identification of university patents related to the SDGs, as they are deemed to be more oriented towards basicness and therefore make less references to practical applications (Trajtenberg et al., 1997). At the same time, the keywords added by the TF-IDF method contain some "noise" that could affect the overall quality of the results obtained.

Future works should consider different approaches to the patent tagging problem, better exploiting the vectorial representations of patent text, in order to understand how patents are written, so to choose the best text representation model. The ideal situation would be to have a *ground truth* about patents related to the SDGs, allowing classification through more sophisticated machine learning methodologies, therefore leading to the creation of a validated dataset. A combination of these techniques, as well as using other patent features, might also allow classification of those patents that do not explicitly mention SDGs related keywords in their text.

Table 3.8: Comparison of results between the two rounds of matching

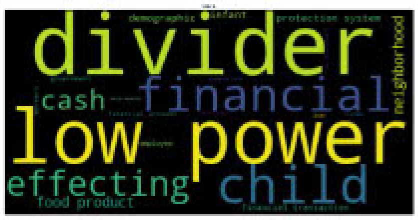
SDG	**No. keywords**	**No. keywords_machted**	**Intersection_TFIDF**	**First_matching**	**TFIDF_keywords_matched**	**Second_matching**	**Total_matches**
SDG1	63	16	6	617	52	66’381	66’998
SDG2	135	47	28	15’055	107	24’312	39’367
SDG3	214	151	107	91’903	258	9’226	101’129
SDG4	156	43	16	1’560	134	44’695	46’255
SDG5	117	13	4	3’218	63	132’603	135’821
SDG6	163	112	82	15’380	152	3’332	18’712
SDG7	161	120	78	75’282	227	10’479	85’761
SDG8	155	44	13	2’600	82	20’694	23’294
SDG9	79	34	10	12’414	56	13’857	26’271
SDG10	115	27	6	668	44	7’488	8’156
SDG11	171	112	65	220’220	168	14’984	235’204
SDG12	150	76	46	13’013	130	18’695	31’708
SDG13	107	36	15	1’647	45	10’779	12’426
SDG14	127	38	18	5’664	96	3’723	8’949
SDG15	179	66	29	3’337	120	6’874	10’211
SDG16	130	41	11	2’241	50	9’248	11’306
Total	2222	976	534	426’863	1’784	348’354	**693’571**

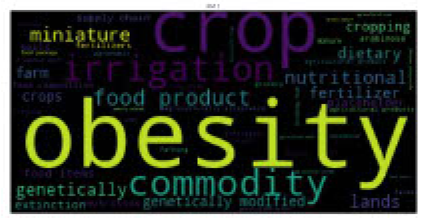
Notes: The second column of the table (No. keywords) represents the number of keywords in the original SDGs lists. The third (No. keywords_matched) is the number of keywords that produced at least one match in patent corpus. The fourth column (Intersection_TFIDF) represents the number of common keywords, among the top 10 identified by the TF-IDF, with the original lists. The fifth column (First_matching) is the number of patents whose text at least matched with 1 keyword of the original list. The sixth column (TFIDF_keywords_matched) is the number of TF-IDF identified extra keywords that produced at least one match in the patent corpus. The seventh column (Second_matching) is the number of patents whose text at least matched with one of the extra-keywords produced by the TF-IDF. Finally, the last column (Total_matches) represents the number of total matches for each of SDG considering both rounds of matching. Note that the total number of first and second matches as well of the total matches does not double count patents assigned to more than one SDG or patents double labelled in both the matchings.

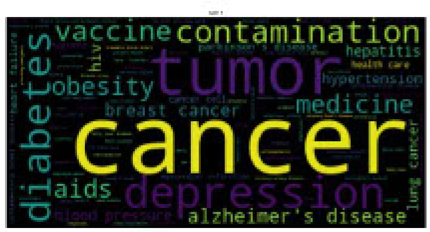
3.6 Appendix for Chapter 3

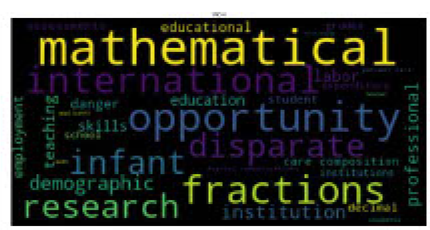
Additional figures and tables

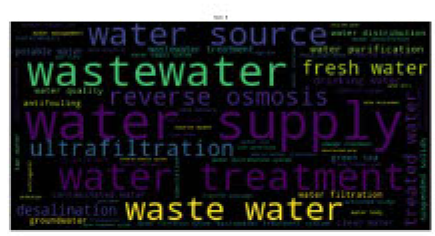
Figure 3.7: Wordclouds for each SDG



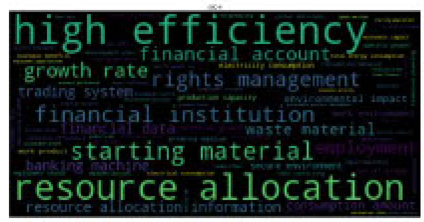


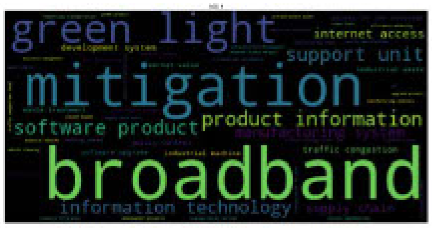




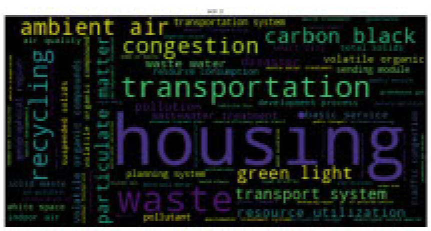


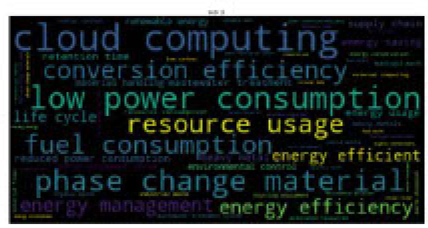

power consumption
fuel cell
photovoltaic
energy storage
solar cell
wind turbine
energy consumption
thermal energy
electric vehicle
solar cells
fuel cells
energy efficiency
lithium battery
lithium ion battery
solar energy
photovoltaic cell
photovoltaic cells
solar power


high efficiency
resource allocation
financial institution
starting material
rights management
financial account
growth rate
trading system
financial data
waste material
banking machine
employment
resource allocation information

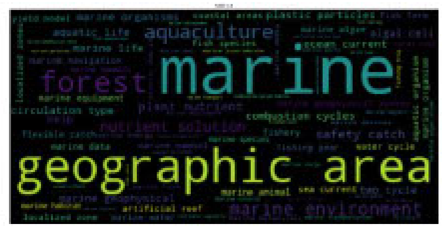

broadband
mitigation
green light
product information
software product
information technology
support unit
internet access
manufacturing system
development system
traffic congestion
supply chain

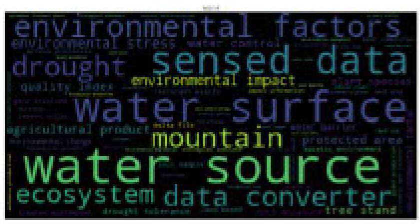

access rights
financial account
sensitive information
health care
insurance policy
social security
financial data
money transfer
insurance coverage


housing
transportation
waste
recycling
particulate matter
ambient air
congestion
carbon black
green light
transport system
resource utilization
transportation system
waste water
volatile organic
pollutant


cloud computing
low power consumption
conversion efficiency
resource usage
fuel consumption
phase change material
energy management
energy efficiency
energy efficient
life cycle
supply chain


management apparatus
environmental data
water level
system management
cleaning composition
nitrous oxide
control strategy
environmental impact
acidification
high altitude
evacuation system
emergency event
sea surface
co2 capture
weather prediction
sea level


marine
geographic area
forest
aquaculture
marine environment
nutrient solution
plant nutrient
aquatic life
marine life
fish species
ocean current
safety catch
circulation type
combustion cycles
artificial reef
marine geophysical
plastic particles
coastal areas
marine organisms
algal cell
marine animal
sea current


environmental factors
sensed data
drought
water surface
water source
mountain
ecosystem
data converter
environmental impact
environmental stress
water control
quality index
agricultural product
protected area
tree stand
drought tolerance


corruption
flow stream
warfare
severity level
rule set
financial institution
financial transactions
identity information
power reduction
rule based
using human
crime
pivot aras
electronic warfare
identity theft
money laundering

	raostirling	family_size	no_inv	bwd_cits	claims
raostirling	1.0000				
family_size	0.0781*	1.0000			
no_inv	0.0319*	0.0319*	1.0000		
bwd_cits	0.0188*	0.1070*	0.0617*	1.0000	
claims	-0.0163*	0.0351*	0.0547*	0.0837*	1.0000

Table 3.9: Pairwise correlation

Notes: Significance level 0.05 or more.

Additional descriptive statistics

Table 3.9 reports the correlation between variables used in our regression analysis.

Additional econometric model

In addition to the models provided, we deem important observing the effect of different kinds of SDGs on technological diversity: the environmental related SDGs (including SDG 6, SDG 7, SDG 11, SDG 13, SDG 14 and SDG 15) (Guo et al., 2020), the "social" related SDGs (including SDG 1, SDG 4, SDG 5, SDG 12, SDG 10, SDG 12, SDG 16) (van der Waal et al., 2021) and the "development" related SDGs (including SDG 2, SDG 3, SDG8 and SDG 9) (WIPO, 2019). Table 3.10 reports the results of the following specifications:

$$\begin{aligned} \Delta_i = \alpha + \beta_1 SDG_i + \beta_2 inventors_i + \beta_3 familySize_i \\ + \beta_4 backCits_i + \beta_5 claims_i + \beta_6 univ_i + IPC.3digit_i + t_i + w_i + \epsilon_i \end{aligned} \tag{3.12}$$

where the variable SDG_i respectively correspond to environmental SDGs (SDG_env), social SDGs (SDG_social) and development SDGs ($SDG\_develop$).

Table 3.10: OLS regression results of the 3 kinds of SDGs on Rao-Stirling with IPC, time and state controls

	(1) Environmental	(2) Social	(3) Development
SDG_env	0.0014***		
	(0.0002)		
family_size	0.0001***	0.0001***	0.0001***
	(0.0000)	(0.0000)	(0.0000)
bwd_cits	-0.0000**	-0.0000**	-0.0000*
	(0.0000)	(0.0000)	(0.0000)
claims	0.0001***	0.0001***	0.0001***
	(0.0000)	(0.0000)	(0.0000)
no_inv	0.0001***	0.0001***	0.0001***
	(0.0000)	(0.0000)	(0.0000)
university	0.0049***	0.0049***	0.0047***
	(0.0003)	(0.0003)	(0.0003)
SDG_social		0.0007***	
		(0.0001)	
SDG_develop			0.0019***
			(0.0002)
_cons	0.1252	0.1253	0.1254
	(0.0874)	(0.0874)	(0.0875)
Observations	1794533	1794533	1794533
Year Dummies	YES	YES	YES
State Dummies	YES	YES	YES
IPC.3digit	YES	YES	YES
R^2	0.167	0.167	0.167

Standard errors in parentheses

* $p < .10$, ** $p < .05$, *** $p < .01$

Notes: The dependent variable is the Rao-Stirling index as defined by Rao (1982). Unit of observation: patent. Grant years: 2006-2020. Heteroskedastic-Robust standard errors in parentheses; ***p<0.01,**p<0.05, *p<0.1

Conclusion

This thesis explored two relevant topics for economics of innovation: patent litigation and the green transition with their implications and interplay.
In the first chapter, we analyzed the relationship between universities and patent litigation, looking for evidence about universities' monetization strategies. Universities, since the Bayh-Dole Act, put a greater emphasis and resources on their technology transfer activities and established TTOs. As non-practicing entities, they also started to enforce their rights in courts, sometimes being accused of looking like patent trolls. Our study is the first to compare the litigation strategies of universities with those of patent trolls. On the one hand, we find that universities, unlike patent trolls, litigate valuable patents, do not do *forum shopping* and do not assert their patents against multiple defendants. However, the picture is quite different when we consider the ICT field, which is also the most targeted by trolls. In this particular field, the characteristics of patents litigated by universities are more similar to those litigated by trolls than to those litigated by other entities. These results shed some clarity on this emerging phenomenon of university patent litigation and calls for further consideration from policy makers.

In the second chapter, considering the increasing attention on the dramatic effects of climate change and the push towards the green transition, we empirically examined the relationship between the regulation-induced prospects for profits in the markets for green technologies and private agents' strategies to reap them, based on patent ligation dynamics in the US. Environmental regulation is deemed to be a trigger for the diffusion of eco-innovations. In this sense, we expect that the regulatory push-pull effect of environmental regulation not only pushes towards the adoption of these technologies, but also stimulate the commitment of private resources to their generation. However, the money flow in this market might attract the interest of actors interested to reap them, such as patent trolls. Hence, we analyze the relationship between patent litigation and the environmental stringency in the US, proxied by green public procurement. We find that, despite the prominence of green technologies, those are not at a higher risk of being litigated. One feasible explanation is that the complexity of green inventions make them more difficult to be litigated, consistently with the literature highlighting the negative relationship between patent breadth and likelihood of litigation. On the other hand, when we consider litigation involving PAE we find that stricter environmental litigation makes them more interested in litigating green technologies. These findings underline the necessity for policy makers to consider this "side effects" of environmental regulation, which in our has case has been proven to attract PAE, thus to increase the cost of adoption of green technologies.

In the third chapter, we studied the relationship between SDGs related innovation and technological diversity, proxied by the Rao-Stirling diversity index. In particular, we focused on patent production from American universities. This because although the link between diverse technologies and complex challenges has already been made explicit in green innovation and SDGs related literature, little attention has been devoted to the diversity of SDGs related innovation and on the contribution of universities, which are deemed to be fundamental to the achievement of the SDGs through a mix of intedisciplinary education, research and innovation. The results show that, consistently with the literaure on green patents, SDGs related patents are more diverse compared to other patents across almost all the technological fields. However, when we look at university patents specifically, only for few SDGs there is a diversity premium. This could suggest that universities often benefit from an interdisciplinary environment whose potential has not yet been fully exploited in the production of sustainable development technologies. These findings could have important policy implications by suggesting to policy makers that, in order to promote technologies for sustainable development, it is necessary to establish partnerships and networks that look beyond the culture of strict disciplinarity and benefit from a wide range of skills and competencies that can be combined.

Bibliography

Abbott, F. M. (2002), ‘Wto trips agreement and its implications for access to medicines in developing countries’, *Study Paper 2a, United Kingdom Commission on Intellectual Property Rights* .

Abrahamson, E. and Baumard, P. (2008), What lies behind organizational façades and how organizational façades lie: An untold story of organizational decision making, *in* ‘The Oxford handbook of organizational decision making’.

Adams, J. D. (1990), ‘Fundamental stocks of knowledge and productivity growth’, *Journal of political economy* 98(4), 673–702.

Ahuja, G. and Morris Lampert, C. (2001), ‘Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions’, *Strategic management journal* 22(6-7), 521–543.

Albareda-Tiana, S., Vidal-Raméntol, S. and Fernández-Morilla, M. (2018), ‘Implementing the sustainable development goals at university level’, *International Journal of Sustainability in Higher Education* .

Allison, J. R., Lemley, M. A., Moore, K. A. and Trunkey, R. D. (2003), ‘Valuable patents’, *Geo. Lj* 92, 435.

Allison, J. R., Lemley, M. A. and Schwartz, D. L. (2013), ‘Understanding the realities of modern patent litigation’, *Tex. L. Rev.* 92, 1769.

Allison, J. R., Lemley, M. A. and Schwartz, D. L. (2017), ‘How often do non-practicing entities win patent suits?’, *Berkeley Technology Law Journal* 32(1), 237–310.

Allison, J. R., Lemley, M. A. and Walker, J. (2009), ‘Extreme value or trolls on top-the characteristics of the most-litigated patents’, *U. Pa. L. Rev.* 158, 1.

Allison, J. R., Lemley, M. A. and Walker, J. (2010), ‘Patent quality and settlement among repeat patent litigants’, *Geo. LJ* 99, 677.

Alstott, J., Triulzi, G., Yan, B. and Luo, J. (2017), ‘Mapping technology space by normalizing patent networks’, *Scientometrics* 110(1), 443–479.

Álvarez, I., Etxeberria, P., Alberdi, E., Pérez-Acebo, H., Eguia, I. and García, M. J. (2021), 'Sustainable civil engineering: Incorporating sustainable development goals in higher education curricula', *Sustainability* 13(16), 8967.

Annan-Diab, F. and Molinari, C. (2017), 'Interdisciplinarity: Practical approach to advancing education for sustainability and for the sustainable development goals', *The International Journal of Management Education* 15(2), 73–83.

Arrow, K. (1962), Economic welfare and the allocation of resources to invention, rate and direction of inventive activity in richard r. nelson (ed.), the rate and direction of inventive activity: Economic and social factors, NBER (National Bureau of Economic Research) Conference Series, Princeton . . . .

Arthur, W. B. (2007), 'The structure of invention', *Research policy* 36(2), 274–287.

Ascione, G. S., Ciucci, L., Detotto, C. and Sterzi, V. (2021), 'Universities and patent trolls: an empirical study of university patent infringement litigation in the united states', *ISSI Paper Proceedings* .

Audia, P. G. and Goncalo, J. A. (2007), 'Past success and creativity over time: A study of inventors in the hard disk drive industry', *Management science* 53(1), 1–15.

Ayres, I. and Ouellette, L. L. (2016), 'A market test for bayh-dole patents', *Cornell L. Rev.* 102, 271.

Ayres, R. U. (2008), 'Sustainability economics: Where do we stand?', *Ecological economics* 67(2), 281–310.

Bannerman, S. (2020), 'The world intellectual property organization and the sustainable development agenda', *Futures* 122, 102586.

Barbieri, N., Ghisetti, C., Gilli, M., Marin, G. and Nicolli, F. (2016), 'A survey of the literature on environmental innovation based on main path analysis', *Journal of Economic Surveys* 30(3), 596–623.

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

Beethika Khan, C. R. and Okrent, A. (2020), 'The state of u.s. science and engineering 2020'.

Bercovitz, J. and Feldman, M. (2006), 'Entpreprenerial universities and technology transfer: A conceptual framework for understanding knowledge-based economic development', *The journal of technology transfer* 31(1), 175–188.

Bessen, J. (2005), 'Patents and the diffusion of technical information', *Economics Letters* 86(1), 121–128.

Bessen, J. E. and Meurer, M. J. (2008), 'The private costs of patent litigation', *Boston University School of Law Working Paper* (07-08).

Bessen, J. and Meurer, M. J. (2013*a*), 'The direct costs from npe disputes', *Cornell L. Rev.* 99, 387.

Bessen, J. and Meurer, M. J. (2013*b*), 'The patent litigation explosion', *Loy. U. Chi. LJ* 45, 401.

Biancani, S., Dahlander, L., McFarland, D. A. and Smith, S. (2018), 'Superstars in the making? the broad effects of interdisciplinary centers', *Research Policy* 47(3), 543–557.

Blakeney, M. (2009), *Intellectual property rights and food security*, Cabi.

Blasco, N., Brusca, I. and Labrador, M. (2020), 'Drivers for universities' contribution to the sustainable development goals: An analysis of spanish public universities', *Sustainability* 13(1), 89.

Blau, P. M. (1977), *Inequality and heterogeneity: A primitive theory of social structure*, Vol. 7, Free Press New York.

Blohmke, J. (2014), 'Technology complexity, technology transfer mechanisms and sustainable development', *Energy for Sustainable Development* 23, 237–246.

Boldrin, M. and Levine, D. (2002), 'The case against intellectual property', *American Economic Review* 92(2), 209–212.

Bonino, D., Ciaramella, A. and Corno, F. (2010), 'Review of the state-of-the-art in patent information and forthcoming evolutions in intelligent patent informatics', *World Patent Information* 32(1), 30–38.

Bottazzi, G. and Pirino, D. (2010), 'Measuring industry relatedness and corporate coherence', *Available at SSRN 1831479* .

Carter-Johnson, J. (2020), University technology transfer structure and intellectual property policies, *in* 'Research Handbook on Intellectual Property and Technology Transfer', Edward Elgar Publishing.

Caviggioli, F., De Marco, A., Montobbio, F. and Ughetto, E. (2020), 'The licensing and selling of inventions by us universities', *Technological Forecasting and Social Change* 159, 120189.

CEA (2016), The patent litigation landscape: Recent research and developments, Technical report.

Ceulemans, K., Lozano, R. and Alonso-Almeida, M. D. M. (2015), 'Sustainability reporting in higher education: Interconnecting the reporting process and organisational change management for sustainability', *Sustainability* 7(7), 8881–8903.

Chen, J., Chen, J., Zhao, S., Zhang, Y. and Tang, J. (2020), 'Exploiting word embedding for heterogeneous topic model towards patent recommendation', *Scientometrics* 125(3), 2091–2108.

Cheng, W., Appolloni, A., D'Amato, A. and Zhu, Q. (2018), 'Green public procurement, missing concepts and future trends–a critical review', *Journal of Cleaner Production* 176, 770–784.

Chien, C. (2013), 'Startups and patent trolls', *Stan. Tech. L. Rev.* 17, 461.

Chien, C. V. (2010), 'From arms race to marketplace: the complex patent ecosystem and its implications for the patent system', *Hastings Lj* 62, 297.

Chien, C. V. (2011), 'Predicting patent litigation', *Tex. L. Rev.* 90, 283.

Christensen, J. M. and Olhoff, A. (2019), 'Emissions gap report 2019', *United Nations Environment Programme (UNEP): Gigiri Nairobi, Kenya* .

Cohen, L., Gurun, U. G. and Kominers, S. D. (2016), 'The growing problem of patent trolling', *Science* 352(6285), 521–522.

Cohen, L., Gurun, U. G. and Kominers, S. D. (2019), 'Patent trolls: Evidence from targeted firms', *Management Science* 65(12), 5461–5486.

Cohen, W. M. and Levinthal, D. A. (1989), 'Innovation and learning: the two faces of r & d', *The economic journal* 99(397), 569–596.

Collinsworth, B. and Crager, S. E. (2014), 'Should academic therapeutic patents go to the highest bidder?'.

Colombelli, A., Ghisetti, C. and Quatraro, F. (2020), 'Green technologies and firms' market value: a micro-econometric analysis of european firms', *Industrial and Corporate Change* 29(3), 855–875.

Colombelli, A., Krafft, J. and Quatraro, F. (2021), 'Firms' growth, green gazelles and eco-innovation: Evidence from a sample of european firms', *Small Business Economics* 56(4), 1721–1738.

Cordova, A. K. and Feldman, R. (2015), 'Universities and patent demands', *Journal of Law and the Biosciences* 2(3), 717–721.

Cordova, M. F. and Celone, A. (2019), 'Sdgs and innovation in the business context literature review', *Sustainability* 11(24), 7043.

Costantini, V., Crespi, F., Martini, C. and Pennacchio, L. (2015), 'Demand-pull and technology-push public support for eco-innovation: The case of the biofuels sector', *Research Policy* 44(3), 577–595.

Council, N. R., Merrill, S. A., Mazza, A.-M. et al. (2011), *Managing university intellectual property in the public interest*, National Academies Press.

Coursey, C. C. (2009), 'Battling the patent troll: tips for defending patent infringement claims by non-manufacturing patentees', *Am. J. Trial Advoc.* 33, 237.

Crespi, F., Ghisetti, C. and Quatraro, F. (2015), 'Environmental and innovation policies for the evolution of green technologies: A survey and a test', *Eurasian Business Review* 5(2), 343–370.

Crespi, F. and Quatraro, F. (2013), ‘Systemic technology policies: Issues and instruments’.

Crespi, F. and Quatraro, F. (2015), *The economics of knowledge, innovation and systemic technology policy*, Vol. 64, Routledge London and New York.

Curtin, R. S. (2014), ‘Slapping patent trolls: What anti-trolling legislation can learn from the anti-slapp movement’, *Stan. Tech. L. Rev.* 18, 39.

De Marchi, V. (2012), ‘Environmental innovation and r&d cooperation: Empirical evidence from spanish manufacturing firms’, *Research policy* 41(3), 614–623.

De Marchi, V. and Grandinetti, R. (2013), ‘Knowledge strategies for environmental innovations: the case of italian manufacturing firms’, *Journal of knowledge management* .

De Rassenfosse, G., Kozak, J. and Seliger, F. (2019), ‘Geocoding of worldwide patent data’, *Scientific data* 6(1), 1–15.

Dechezleprêtre, A., Martin, R. and Mohnen, M. (2013), *Knowledge spillovers from clean and dirty technologies: A patent citation analysis*, Grantham Research Institute on Climate Change and the Environment.

del Río González, P. (2009), ‘The empirical analysis of the determinants for environmental technological change: A research agenda’, *Ecological economics* 68(3), 861–878.

Deuten, J. (2003), ‘Cosmopolitanising technologies. a study of four emerging technological regimes’.

Drivas, K., Lei, Z. and Wright, B. D. (2017), ‘Academic patent licenses: Roadblocks or signposts for nonlicensee cumulative innovation?’, *Journal of Economic Behavior & Organization* 137, 282–303.

Eisenberg, R. S. and Cook-Deegan, R. (2018), ‘Universities: The fallen angels of bayh-dole?’, *Daedalus* 147(4), 76–89.

El-Jardali, F., Ataya, N. and Fadlallah, R. (2018), ‘Changing roles of universities in the era of sdgs: rising up to the global challenge through institutionalising partnerships with governments and communities’, *Health research policy and systems* 16(1), 1–5.

Engelsman, E. C. and van Raan, A. F. (1994), ‘A patent-based cartography of technology’, *Research Policy* 23(1), 1–26.

Etzkowitz, H. (1990), The second academic revolution: The role of the research university in economic development, *in* ‘The research system in transition’, Springer, pp. 109–124.

Etzkowitz, H. (2001), ‘The second academic revolution and the rise of entrepreneurial science’, *IEEE Technology and Society Magazine* 20(2), 18–29.

Federal Trade Commission, U. S. (2016), 'Patent assertion entity activity: An ftc study', *Washington, DC: US Federal Trade Commission* .

Feldman, R. and Ewing, T. (2012), 'The giants among us', *Stanford Technology Law Review* 1.

Feldman, R. and Lemley, M. A. (2015), 'Do patent licensing demands mean innovation', *Iowa L. Rev.* 101, 137.

Feng, J. and Jaravel, X. (2016), Who feeds the trolls? patent trolls and the patent examination process, Technical report, Working Paper.

Feng, J. and Jaravel, X. (2020), 'Crafting intellectual property rights: Implications for patent assertion entities, litigation, and innovation', *American Economic Journal: Applied Economics* 12(1), 140–81.

Feng, S. (2020), 'The proximity of ideas: An analysis of patent text using machine learning', *PloS one* 15(7), e0234880.

Firpo, T. and Mireles, M. S. (2018), 'Monitoring behavior: Universities, nonprofits, patents, and litigation', *SMUL Rev.* 71, 505.

Firpo, T. and Mireles, M. S. (2020), Currents and crosscurrents in litigation of university and nonprofit related patents: is there a coming wave of patent litigation involving those patents?, *in* 'Research Handbook on Intellectual Property and Technology Transfer', Edward Elgar Publishing.

Fischer, T. and Henkel, J. (2012), 'Patent trolls on markets for technology–an empirical analysis of npes' patent acquisitions', *Research Policy* 41(9), 1519–1533.

Fusco, S., Lissoni, F., Martinez, C. and Sterzi, V. (2019), 'Monetization strategies of university patents through paes: an analysis of us patent transfers', *Available at SSRN 3410086* .

Fusillo, F., Quatraro, F. and Santhià, C. (2021), The geography of circular economy technologies in europe: evolutionary patterns and technological convergence, *in* 'Research Handbook of Innovation for a Circular Economy', Edward Elgar Publishing.

Fusillo, F. et al. (2020), Are green inventions really more complex? evidence from european patents, Technical report, University of Turin.

Gagliardi, L., Marin, G. and Miriello, C. (2016), 'The greener the better? job creation effects of environmentally-friendly technological change', *Industrial and Corporate Change* 25(5), 779–807.

Ghisetti, C. and Quatraro, F. (2013), 'Beyond inducement in climate change: Does environmental performance spur environmental technologies? a regional analysis of cross-sectoral differences', *Ecological Economics* 96, 99–113.

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

Ghosh, S. (2016), ‘Are universities special’, *Akron L. Rev.* 49, 671.

Günther, E. and Scheibe, L. (2006), ‘The hurdle analysis. a self-evaluation tool for municipalities to identify, analyse and overcome hurdles to green procurement’, *Corporate Social Responsibility and Environmental Management* 13(2), 61–77.

Guo, M., Nowakowska-Grunt, J., Gorbanyov, V. and Egorova, M. (2020), ‘Green technology and sustainable development: Assessment and green growth frameworks’, *Sustainability* 12(16), 6571.

Hagiu, A. and Yoffie, D. B. (2013), ‘The new patent intermediaries: platforms, defensive aggregators, and super-aggregators’, *Journal of Economic Perspectives* 27(1), 45–66.

Hajikhani, A. and Suominen, A. (2021), ‘The interrelation of sustainable development goals in publications and patents: A machine learning approach’.

Hargadon, A. and Sutton, R. I. (1997), ‘Technology brokering and innovation in a product development firm’, *Administrative science quarterly* pp. 716–749.

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

Heinzl, J., Kor, A.-L., Orange, G. and Kaufmann, H. R. (2013), ‘Technology transfer model for austrian higher education institutions’, *The Journal of Technology Transfer* 38(5), 607–640.

Henderson, R., Jaffe, A. B. and Trajtenberg, M. (1998), ‘Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988’, *Review of Economics and statistics* 80(1), 119–127.

Henry, C. and Stiglitz, J. E. (2010), ‘Intellectual property, dissemination of innovation and sustainable development’, *Global Policy* 1(3), 237–251.

Hernandez-Aguilera, J. N., Anderson, W., Bridges, A. L., Fernandez, M. P., Hansen, W. D., Maurer, M. L., Ilboudo Nebie, E. K. and Stock, A. (2021), ‘Supporting interdisciplinary careers for sustainability’, *Nature Sustainability* 4(5), 374–375.

Hoppmann, J., Peters, M., Schneider, M. and Hoffmann, V. H. (2013), ‘The two faces of market support—how deployment policies affect technological exploration and exploitation in the solar photovoltaic industry’, *Research policy* 42(4), 989–1003.

IAM Magasine (2019), Issue 94 March/April.

Idris, K. (2003), ‘A power tool for economic growth’, *WIPO publication* 888.

Jaffe, A. B. and Lerner, J. (2011), Innovation and its discontents, *in* 'Innovation and Its Discontents', Princeton University Press.

Jaffe, A. B. and Trajtenberg, M. (2002), *Patents, citations, and innovations: A window on the knowledge economy*, MIT press.

Jensen, R. and Thursby, M. (2001), 'Proofs and prototypes for sale: The licensing of university inventions', *American Economic Review* 91(1), 240–259.

Johnstone, N., Haščič, I., Poirier, J., Hemar, M. and Michel, C. (2012), 'Environmental policy stringency and technological innovation: evidence from survey data and patent counts', *Applied Economics* 44(17), 2157–2170.

Jones, P., Selby, D. and Sterling, S. (2010), 'More than the sum of their parts? interdisciplinarity and sustainability', *Sustainability education: Perspectives and practice across higher education* pp. 17–38.

Kemp, R. (2010), 'Eco-innovation: definition, measurement and open research issues', *Economia politica* 27(3), 397–420.

Kesan, J. P. (2008), 'Transferring innovation', *Fordham L. Rev.* 77, 2169.

Kestin, T., van den Belt, M., Denby, L., Ross, K., Thwaites, J. and Hawkes, M. (2017), 'Getting started with the sdgs in universities: A guide for universities, higher education institutions, and the academic sector'.

Khachatryan, D. and Muehlmann, B. (2019), 'Measuring technological breadth and depth of patent documents using rao's quadratic entropy', *Journal of Applied Statistics* .

Kiebzak, S., Rafert, G. and Tucker, C. E. (2016), 'The effect of patent litigation and patent assertion entities on entrepreneurial activity', *Research Policy* 45(1), 218–231.

Kingston, W. (2001), 'Innovation needs patents reform', *Research Policy* 30(3), 403–423.

Kopnina, H. (2020), 'Education for the future? critical evaluation of education for sustainable development goals', *The Journal of Environmental Education* 51(4), 280–291.

Körfgen, A., Förster, K., Glatz, I., Maier, S., Becsi, B., Meyer, A., Kromp-Kolb, H. and Stötter, J. (2018), 'It'sa hit! mapping austrian research contributions to the sustainable development goals', *Sustainability* 10(9), 3295.

Kostoska, O. and Kocarev, L. (2019), 'A novel ict framework for sustainable development goals', *Sustainability* 11(7), 1961.

Lam, J. C., Walker, R. M. and Hills, P. (2014), ‘Interdisciplinarity in sustainability studies: a review’, *Sustainable Development* 22(3), 158–176.

Lanjouw, J. O. and Schankerman, M. (2001), ‘Characteristics of patent litigation: a window on competition’, *RAND journal of economics* pp. 129–151.

Leal Filho, W., Shiel, C., Paço, A., Mifsud, M., Ávila, L. V., Brandli, L. L., Molthan-Hill, P., Pace, P., Azeiteiro, U. M., Vargas, V. R. et al. (2019), ‘Sustainable development goals and sustainability teaching at universities: Falling behind or getting ahead of the pack?’, *Journal of Cleaner Production* 232, 285–294.

Ledford, H. (2013), ‘Universities struggle to make patents pay’, *Nature* 501(7468), 471–472.

Lemley, M. A. (2007), ‘Are universities patent trolls’, *Fordham Intell. Prop. Media & Ent. LJ* 18, 611.

Lemley, M. A. and Feldman, R. (2016), ‘Patent licensing, technology transfer, and innovation’, *American Economic Review* 106(5), 188–92.

Lemley, M. A. and Feldman, R. (2020), Is patent enforcement efficient?, *in* ‘Research Handbook on Intellectual Property and Technology Transfer’, Edward Elgar Publishing.

Lemley, M. A. and Shapiro, C. (2005), ‘Probabilistic patents’, *Journal of Economic Perspectives* 19(2), 75–98.

Leoncini, R., Marzucchi, A., Montresor, S., Rentocchini, F. and Rizzo, U. (2019), “better late than never’: the interplay between green technology and age for firm growth’, *Small Business Economics* 52(4), 891–904.

Lerner, J. (1994), ‘The importance of patent scope: an empirical analysis’, *The RAND Journal of Economics* pp. 319–333.

Leydesdorff, L. (2018), ‘Diversity and interdisciplinarity: how can one distinguish and recombine disparity, variety, and balance?’, *Scientometrics* 116(3), 2113–2121.

Leydesdorff, L., Wagner, C. S. and Bornmann, L. (2019), ‘Interdisciplinarity as diversity in citation patterns among journals: Rao-stirling diversity, relative variety, and the gini coefficient’, *Journal of Informetrics* 13(1), 255–269.

Li, S., Hu, J., Cui, Y. and Hu, J. (2018), ‘Deeppatent: patent classification with convolutional neural networks and word embedding’, *Scientometrics* 117(2), 721–744.

Liang, M. (2010), ‘The aftermath of ts tech: The end of forum shopping in patent litigation and implications for non-practicing entities’, *Tex. Intell. Prop. LJ* 19, 29.

Love, B. J. (2013), 'Do university patents pay off-evidence from a survey of university inventors in computer science and electrical engineering', *Yale JL & Tech.* 16, 285.

Love, B. J., Oliver, E. and Costa, M. (2020), Us patent sales by universities and research institutes, *in* 'Research Handbook on Intellectual Property and Technology Transfer', Edward Elgar Publishing.

Lowe, R. A. (2002), *Invention, innovation, and entrepreneurship: The commercialization of university research by inventor-founded firms*, University of California, Berkeley.

Lozano, R., Lukman, R., Lozano, F. J., Huisingh, D. and Lambrechts, W. (2013), 'Declarations for sustainability in higher education: becoming better leaders, through addressing the university system', *Journal of Cleaner Production* 48, 10–19.

Lu, J. 2012), 'The myths and facts of patent troll and excessive payment: have nonpracticing entities (npes) been overcompensated?', *Business Economics* 47, 234–249.

Marco, A. C., Myers, A., Graham, S. J., D'Agostino, P. and Apple, K. (2015), 'The uspto patent assignment dataset: Descriptions and analysis'.

Marin, G. (2014), 'Do eco-innovations harm productivity growth through crowding out? results of an extended cdm model for italy', *Research Policy* 43(2), 301–317.

Martínez, C. and Sterzi, V. (2019), University patenting and the quest for technology transfer policy models in europe, *in* 'Handbook of Universities and Regional Development', Edward Elgar Publishing.

Masters, R. M. and Weber, B. S. (2009), 'Intellectual property cases and trends to follow', *Intellectual Property & Technology Law Journal* 21(6), 14.

Mazzoleni, R. (2006), 'The effects of university patenting and licensing on downstream r&d investment and social welfare', *The Journal of Technology Transfer* 31(4), 431–441.

Meadows, D. H., Randers, J. and Meadows, D. L. (1972), *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind.*, New York: Universe Books.

Messeni Petruzzelli, A., Maria Dangelico, R., Rotolo, D. and Albino, V. (2011), 'Organizational factors and technological features in the development of green innovations: Evidence from patent analysis', *Innovation* 13(3), 291–310.

Miller, S. P. (2018), 'Who's suing us: Decoding patent plaintiffs since 2000 with the stanford npe litigation dataset', *Stan. Tech. L. Rev.* 21, 235.

Miller, S. P., Aravind, A., Bengfort, B., De La Cerda, C., Dragoni, M., Gibson, K., Itai, A., Johnson, C., Kannappan, D., Kehoe, E. et al. (2017), Introduction to the stanford npe litigation dataset, Technical report, Tech. rep., Stanford Program in Law, Science, and Technology.

Morse, W. C., Nielsen-Pincus, M., Force, J. E. and Wulfhorst, J. (2007), ‘Bridges and barriers to developing and conducting interdisciplinary graduate-student team research’, *Ecology and Society* 12(2).

Morton, F. M. S. and Shapiro, C. (2013), ‘Strategic patent acquisitions’, *Antitrust LJ* 79, 463.

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

Murray, F. and Stern, S. (2007), ‘Do formal intellectual property rights hinder the free flow of scientific knowledge?: An empirical test of the anti-commons hypothesis’, *Journal of Economic Behavior & Organization* 63(4), 648–687.

Nelsen, L. (2002), ‘The role of university technology transfer operations in assuring access to medicines and vaccines in developing countries’, *Yale J. Health Pol’y L. & Ethics* 3, 301.

Nelson, R. R. (1993), *National innovation systems: a comparative analysis*, Oxford University Press on Demand.

Nemet, G. F. (2009), ‘Demand-pull, technology-push, and government-led incentives for non-incremental technical change’, *Research policy* 38(5), 700–709.

Newell, W. H., Wentworth, J. and Sebberson, D. (2001), ‘A theory of interdisciplinary studies’, *Issues in Interdisciplinary Studies* .

Nilsson, M., Griggs, D. and Visbeck, M. (2016), ‘Policy: map the interactions between sustainable development goals’, *Nature* 534(7607), 320–322.

Noh, H., Jo, Y. and Lee, S. (2015), ‘Keyword selection and processing strategy for applying text mining to patent analysis’, *Expert Systems with Applications* 42(9), 4348–4360.

Norberg-Bohm, V. (2000), ‘Creating incentives for environmentally enhancing technological change: lessons from 30 years of us energy technology policy’, *Technological forecasting and social change* 65(2), 125–148.

Oguamanam, C. (2006), ‘Intellectual property rights in plant genetic resources: Farmers’ rights and food security of indigenous and local communities’, *Drake J. Agric. L.* 11, 273.

O’Neill, C. (2006), ‘The battle over blackberry: Patent trolls and information technology’, *Journal of Law, Information and Science* 17, 94–125.

Orsatti, G., Perruchas, F., Consoli, D. and Quatraro, F. (2020), ‘Public procurement, local labor markets and green technological change. evidence from us commuting zones’, *Environmental and Resource Economics* 75(4), 711–739.

Orsatti, G., Quatraro, F. and Pezzoni, M. (2020), 'The antecedents of green technologies: The role of team-level recombinant capabilities', *Research Policy* 49(3), 103919.

Orsatti, G. and Sterzi, V. (2018), 'Do patent assertion entities harm innovation? evidence from patent transfers', *Cahiers du Gretha* (2018-08).

Orsenigo, L., Sterzi, V. et al. (2010), 'Comparative study of the use of patents in different industries', *Knowledge, Internationalization and Technology Studies (KITeS)* 33, 1–31.

Ouellette, L. L. and Weires, R. (2019), 'University patenting: Is private law serving public values?', *Michigan State Law Review, Forthcoming* .

Owens, T. L. (2017), 'Higher education in the sustainable development goals framework', *European Journal of Education* 52(4), 414–420.

Özel, S. Ö. and Pénin, J. (2016), 'Exclusive or open? an economic analysis of university intellectual property patenting and licensing strategies', *Journal of Innovation Economics Management* (3), 133–153.

Papadimitriou, A. (2020), 'Beyond rhetoric: Reinventing the public mission of higher education'.

Papke, L. E. and Wooldridge, J. M. (1996), 'Econometric methods for fractional response variables with an application to 401 (k) plan participation rates', *Journal of applied econometrics* 11(6), 619–632.

Parikka-Alhola, K. (2008), 'Promoting environmentally sound furniture by green public procurement', *Ecological economics* 68(1-2), 472–485.

Pohlmann, T. and Opitz, M. (2013), 'Typology of the patent troll business', *R&D Management* 43(2), 103–120.

Popp, D. and Newell, R. (2012), 'Where does energy r&d come from? examining crowding out from energy r&d', *Energy economics* 34(4), 980–991.

Porter, A., Cohen, A., David Roessner, J. and Perreault, M. (2007), 'Measuring researcher interdisciplinarity', *Scientometrics* 72(1), 117–147.

Porter, A. and Rafols, I. (2009), 'Is science becoming more interdisciplinary? measuring and mapping six research fields over time', *Scientometrics* 81(3), 719–745.

Porter, M. E. and Van der Linde, C. (1995), 'Toward a new conception of the environment-competitiveness relationship', *Journal of economic perspectives* 9(4), 97–118.

Quatraro, F. and Scandura, A. (2019), 'Academic inventors and the antecedents of green technologies. a regional analysis of italian patent data', *Ecological Economics* 156, 247–263.

Rafols, I. and Meyer, M. (2010), 'Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience', *Scientometrics* 82(2), 263–287.

Rao, C. R. (1982), 'Diversity and dissimilarity coefficients: a unified approach', *Theoretical population biology* 21(1), 24–43.

Reimers, N. and Gurevych, I. (2019), 'Sentence-bert: Sentence embeddings using siamese bert-networks', *arXiv preprint arXiv:1908.10084* .

Reitzig, M., Henkel, J. and Heath, C. (2007), 'On sharks, trolls, and their patent prey—unrealistic damage awards and firms' strategies of "being infringed"', *Research Policy* 36(1), 134–154.

Reitzig, M., Henkel, J. and Schneider, F. (2010), 'Collateral damage for r&d manufacturers: how patent sharks operate in markets for technology', *Industrial and Corporate Change* 19(3), 947–967.

Rennings, K. (2000), 'Redefining innovation—eco-innovation research and the contribution from ecological economics', *Ecological economics* 32(2), 319–332.

Rimmer, M. (2014), 'Intellectual property and global warming: fossil fuels and climate justice', *The Sage Handbook of Intellectual Property, London: Sage Publications* pp. 727–753.

Rimmer, M. (2018), 'A submission on intellectual property and the united nations sustainable development goals'.

Risch, M. (2012), 'Patent troll myths', *Seton Hall L. Rev.* 42, 457.

Risch, M. (2015), 'A generation of patent litigation', *San Diego L. Rev.* 52, 67.

Rogers, F. T. (2020), 'Patent text similarity and cross-cultural venture-backed innovation', *Journal of Behavioral and Experimental Finance* 26, 100319.

Rooksby, J. H. (2011), 'University initiation of patent infringement litigation', *John Marshall Review of Intellectual Property Law* 10(4), 623.

Rooksby, J. H. (2012), 'Innovation and litigation: tensions between universities and patents and how to fix them', *Yale JL & Tech.* 15, 312.

Rooksby, J. H. (2013), 'When tigers bare teeth: A qualitative study on university patent enforcement', *Akron L. Rev.* 46, 169.

Rooksby, J. H. (2016), *The branding of the American mind: How universities capture, manage, and monetize intellectual property and why it matters*, JHU Press.

Rothaermel, F. T. and Thursby, M. (2005), 'University–incubator firm knowledge flows: assessing their impact on incubator firm performance', *Research policy* 34(3), 305–320.

Roudsari, A. H., Afshar, J., Lee, S. and Lee, W. (2021), Comparison and analysis of embedding methods for patent documents, *in* '2021 IEEE International Conference on Big Data and Smart Computing (BigComp)', IEEE, pp. 152–155.

Rowe, E. A. (2005), 'The experimental use exception to patent infringement: Do universities deserve special treatment', *Hastings LJ* 57, 921.

RPX (2015), Npe litigation, patent marketplace, and npe cost, Technical report, Available at https://www.rpxcorp.com/reports/.

Ryan Jr, C. J. and Frye, B. L. (2017), 'An empirical study of university patent activity', *NYU J. Intell. Prop. & Ent. L.* 7, 51.

Rylance, R. (2015), 'Grant giving: Global funders to focus on interdisciplinarity', *Nature* 525(7569), 313–315.

Sampat, B. N. (2020), 'Whose drugs are these?', *Issues in Science and Technology* 36(4), 42–48.

Sampat, B. N. (2021), 'The government and pharmaceutical innovation: Looking back and looking ahead', *Journal of Law, Medicine & Ethics* 49(1), 10–18.

Sánchez-Carracedo, F., Ruiz-Morales, J., Valderrama-Hernández, R., Muñoz-Rodríguez, J. M. and Gomera, A. (2021), 'Analysis of the presence of sustainability in higher education degrees of the spanish university system', *Studies in Higher Education* 46(2), 300–317.

Schwartz, D. L., Sichelman, T. M. and Miller, R. (2019), 'Uspto patent number and case code file dataset documentation'.

Shane, S. A. (2004), *Academic entrepreneurship: University spinoffs and wealth creation*, Edward Elgar Publishing.

Shane, S. and Somaya, D. (2007), 'The effects of patent litigation on university licensing efforts', *Journal of Economic Behavior & Organization* 63(4), 739–755.

Shrestha, S. K. (2010), 'Trolls or market-makers-an empirical analysis of nonpracticing entities', *Colum. L. Rev.* 110, 114.

Smeets, R. (2014), 'Does patent litigation reduce corporate r&d? an analysis of us public firms', *An Analysis of US Public Firms (April 28, 2014)* .

Squicciarini, M., Dernis, H. and Criscuolo, C. (2013), 'Measuring patent quality: Indicators of technological and economic value'.

Sterzi, V. (2021), 'Patent assertion entities and patent ownership transparency: strategic recording of patent transactions at the uspto', *Journal of Competition Law & Economics* 17(4), 978–1006.

Sterzi, V., Rameshkoumar, J.-P. and Van Der Pol, J. (2021*a*), 'Non-practicing entities and transparency in patent ownership in europe', *Technological Forecasting and Social Change (forthcoming)* .

Sterzi, V., Rameshkoumar, J.-P. and Van Der Pol, J. (2021*b*), 'Non-practicing entities and transparency of patent ownership in europe: the case of uk dormant companies', *Technological Forecasting and Social Change* 172, 121069.

Stirling, A. (2007), 'A general framework for analysing diversity in science, technology and society', *Journal of the Royal Society Interface* 4(15), 707–719.

Tarrant, S. P. and Thiele, L. P. (2017), 'Enhancing and promoting interdisciplinarity in higher education', *Journal of Environmental Studies and Sciences* 7(2), 355–360.

Teece, D. J., Rumelt, R., Dosi, G. and Winter, S. (1994), 'Understanding corporate coherence: Theory and evidence', *Journal of economic behavior & organization* 23(1), 1–30.

Testa, F., Iraldo, F., Frey, M. and Daddi, T. (2012), 'What factors influence the uptake of gpp (green public procurement) practices? new evidence from an italian survey', *Ecological Economics* 82, 88–96.

Thomas, I. (2016), Challenges for implementation of education for sustainable development in higher education institutions, *in* 'Routledge Handbook of Higher Education for Sustainable Development', number 4, RMIT University, pp. 56–71.

Thompson, N. C., Ziedonis, A. A. and Mowery, D. C. (2018), 'University licensing and the flow of scientific knowledge', *Research Policy* 47(6), 1060–1069.

Trajtenberg, M., Henderson, R. and Jaffe, A. (1997), 'University versus corporate patents: A window on the basicness of invention', *Economics of Innovation and new technology* 5(1), 19–50.

Tseng, Y.-H., Lin, C.-J. and Lin, Y.-I. (2007), 'Text mining techniques for patent analysis', *Information processing & management* 43(5), 1216–1247.

UN (2020), Guidebook for the preparation of science, technology and innovation (sti) for sdgs roadmaps, Technical report.

US-FTC (2016), 'Patent assertion entity activity: An ftc study'.

Usui, S., Palmes, P., Nagata, K., Taniguchi, T. and Ueda, N. (2007), 'Keyword extraction, ranking, and organization for the neuroinformatics platform', *Biosystems* 88(3), 334–342.

Valdivia, W. D. (2013), 'University start-ups: Critical for improving technology transfer', *Center for Technology Innovation at Brookings. Washington, DC: Brookings Institution* .

Valdivia, W. D. (2015), 'Patent infringement suits have a reputational cost for universities'.
URL: *https://www.brookings .edu/blog/techtank/2015/11/10/patent-infringement-suits-have-a-reputational-cost-for -universities-2/;*

van der Waal, J. W., Thijssens, T. and Maas, K. (2021), 'The innovative contribution of multinational enterprises to the sustainable development goals', *Journal of Cleaner Production* 285, 125319.

Van Looy, B., Du Plessis, M. and Magerman, T. (2006), 'Data production methods for hamonized patent statistics: Patentee sector allocation'.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M. and Fuso Nerini, F. (2020), 'The role of artificial intelligence in achieving the sustainable development goals', *Nature communications* 11(1), 1–10.

Walz, R., Pfaff, M., Marscheider-Weidemann, F. and Glöser-Chahoud, S. (2017), 'Innovations for reaching the green sustainable development goals–where will they come from?', *International Economics and Economic Policy* 14(3), 449–480.

Wang, J., Thijs, B. and Glänzel, W. (2015), 'Interdisciplinarity and impact: Distinct effects of variety, balance, and disparity', *PloS one* 10(5), e0127298.

Weitzman, M. L. (1998), 'Recombinant growth', *The Quarterly Journal of Economics* 113(2), 331–360.

Winickoff, D. E. (2013), 'Private assets, public mission: The politics of technology transfer and the new american university', *Jurimetrics* pp. 1–42.

WIPO (2012), 'Ipc green inventory'.

WIPO (2018), 'Report on wipo's contribution to the implementation of the sustainable development goals and its associated targets'.

WIPO (2019), 'Report on wipo's contribution to the implementation of the sustainable development goals and its associated targets'.

Xie, Z. and Miyazaki, K. (2013), 'Evaluating the effectiveness of keyword search strategy for patent identification', *World Patent Information* 35(1), 20–30.

Yan, B. and Luo, J. (2017), 'Measuring technological distance for patent mapping', *Journal of the Association for Information Science and Technology* 68(2), 423–437.

Yeh, B. T. (2012), 'An overview of the" patent trolls" debate'.

Zeppini, P. and van Den Bergh, J. C. (2011), 'Competing recombinant technologies for environmental innovation: extending arthur's model of lock-in', *Industry and Innovation* 18(03), 317–334.

Zhu, K., Malhotra, S. and Li, Y. (2022), 'Technological diversity of patent applications and decision pendency', *Research Policy* 51(1), 104364.