Department of Economics and Statistics "Cognetti de Martiis"
Campus Luigi Einaudi, Lungo Dora Siena 100/A, 10153 Torino (Italy)
www.est.unito.it

Dipartimento
Economia e Statistica
Cognetti de Martiis

Working Paper Series

15/20

ARE GREEN INVENTIONS REALLY MORE COMPLEX? EVIDENCE FROM EUROPEAN PATENTS

FABRIZIO FUSILLO

Bureau of Research on Innovation, Complexity and Knowledge

The Department of Economics and Statistics "Cognetti de Martiis" publishes research papers authored by members and guests of the Department and of its research centers. ISSN: 2039-4004

Are Green Inventions really more complex? Evidence from European Patents

Fabrizio Fusillo[a,b]

[a]Department of Economics and Statistics "Cognetti De Martiis", University of Turin
[b]BRICK, Collegio Carlo Alberto

Abstract

A large body of existing literature extensively studied the economic determinants and effects of environmental innovations. However, only a few studies analyzed the specific features of green technologies in the early phases of the invention process. The aim of this paper is to investigate knowledge recombination patterns in the green domain. The focus is on identifying whether and how different bodies of technology are combined and integrated. Exploiting a large sample of European patent data, from 1980 to 2012, the paper investigates the degree of diversity in the knowledge sources and the generation phase of green inventions. Using the Integration Score as an index of technological diversity we compare the recombinant features of Green Technologies with a control sample of "Traditional Technologies", accurately drawn from the universe of all patent applications. Empirical results suggest that, after controlling for a number of typical characteristics which may affect diversity, Green Technologies systematically show a higher degree of diversity when compared to non-green ones.

Keywords: Green technologies, environmental innovation, knowledge recombination, patents, diversity, knowledge space
JEL Classification codes: O31, Q55

1 Introduction

During the last few decades, we have witnessed increasing attention towards the economic analysis of environmental issues. In the transition toward a greener economy, experts underlined the fundamental role of innovation activities to address climate change and environmental challenges. Green technologies are increasingly becoming a major source of technical change since the so-called "green" technological change is seen as a way to reduce environmental pressure and restore production efficiency and competitiveness.

The growing literature dealing with the determinants of environmental innovations identifies regulation as a key driver in the generation and adoption of Green Technologies (henceforth GTs) (Porter and Linde 2011; Carrión-Flores and Innes 2010; Carrión-Flores et al. 2013). Known as "double externality" problem, a major economic issue in environmental innovations is related to the two types of positive externalities they produce: knowledge externalities in the research and innovation phase, leading to market failures and usual sub-optimal investments in R&D; externalities in the adoption phase due to the improvement of environmental quality, making their diffusion socially desirable. In this context, stringent regulatory frameworks, stimulating the adoption of GTs, contributes to the creation of new market niches for GTs, thus, providing incentives for advancements in these fields (Arimura et al. 2007; Brunnermeier and Cohen 2003; Johnstone et al. 2010; Lanoie et al. 2011; Popp 2003; Ghisetti and Quatraro 2015). Therefore, a wide body of literature focuses on the importance of environmental policy and on understanding the innovation and diffusion stages of GTs. Less attention has instead been dedicated to the study of the antecedents of green inventions (Del Río González 2009) and, more specifically, on a direct comparison between GTs and other technologies.

The aim of this work is to fill these gaps, providing new evidence on the recombination patterns of GTs in the early phases of the invention process. The paper contributes to and extends existing literature in several ways. First, we investigate the dynamics underlying the generation of GTs, focusing on how and to what extent green invention make use of more diversified knowledge sources and recombine different pieces of knowledge. In the recombinant knowledge framework, considerable efforts have been dedicated to the study of the cognitive dimension of knowledge recombination – e.g. at the inventors level – and on the characterization of the Knowledge Base of different industries (Breschi et al. 2003; Corrocher et al. 2007; Nesta and Saviotti 2005). Only recently scholars brought to the core of academic debate the dynamics behind green inventions, showing that GTs tend to arise from the recombination of a wide array of technological domains (Dechezleprêtre et al.

2015), often benefiting from the hybridization of both green and non-green technologies (Orsatti et al. 2020; Quatraro and Scandura 2019; Zeppini and Den Bergh 2011).

The second contribution of the paper is related to the understanding of the specific features of GTs with respect to "traditional" or non-green technologies. Existing literature provides evidence that, given their systemic and multi-purpose nature (Ghisetti et al. 2015), the development of green technologies and new green products requires a broader range of skills and competences, which can often be far from the traditional knowledge bases and capabilities (De Marchi 2012; Fusillo et al. 2019). Recent studies, using patent data, found that green patents are more original and radical than non-green patents, and generate larger knowledge spillover then their dirty counterpart (Dechezleprêtre et al. 2015; Barbieri et al. 2020). However, at the invention-technology level, systematic evidence on the intrinsic higher technological diversity in green recombinations compared to non-green is still relatively scant.

Relying on the information contained in patent documents registered at the European Patent Office, this paper explores the evolution of the technological diversity in the green technology domain, during the period 1980-2012. Exploiting the CPC classification of technologies, we construct an indicator of diversity of the knowledge sources (*search phase*) starting from the co-occurrence of technological classes in the *patent citations*; we instead rely on the co-classification of classes in the focal patents to calculate the degree of diversity in the *production phase* (knowledge generation). Hence, we compare the recombinant technological diversification of green technologies with a sample of non-green technologies, selected through propensity score matching, to assess whether and to what extent green technologies draw upon and recombine more diversified technological knowledge.

Results suggest that green inventions source from a much more diversified set of technological knowledge compared to inventions with similar characteristics. Moreover, the generation of green technologies involves the recombination of dispersed pieces of knowledge which are often distant from each other in the technological space.

The remaining of the paper is organized as follow. Section 2 reviews the relevant literature. Section 3 describes the data sources, the methodology and the empirical analysis. Results are presented in Section 4. Section 5 concludes and derives some policy implications.

2 Theoretical framework

2.1 Recombination and green technologies

The extensive literature on technological change provides a deep understanding of the processes of technological diffusion, commercialization and the impact on organizations and economic growth. Aiming at understanding the underlying patterns of technological development (Archibugi and Pianta 1994), recent studies, drawing from evolutionary theory, proposed a view of the inventive activity as the outcome of knowledge recombination processes (Nelson and Winter 1982; Schumpeter 1934). In this view, new technologies originate from the recombination of existing and/or new pieces of knowledge, highly dispersed in the technological space. Thus, given the cumulative and path-dependent nature of technology, the recombinant knowledge approach recognizes the importance of studying the structure of the search space and their impact on the invention process (Fleming and Sorenson 2001; Olsson 2000).

Empirical evidence on specific industries shows that the development pattern of technologies and the characteristic of their knowledge sources may greatly vary within and across different sectors and technological fields. For example, Corrocher et al. (2007) found that ICT inventions may be distinguished into two main groups, one showing a highly diversified knowledge base and high growth in patenting activity, while the other shows a lower rate of growth and less diversified knowledge sources. Krafft et al. (2011) investigated the dynamics of knowledge generation in the biotechnology sector, analyzing the structure of its knowledge base.

Following the momentum gained recently by the literature on the determinants of environmental innovations, several studies applied the recombinant framework to the understanding of the development pattern in GTs. This strand of research defines green innovations as complex and sophisticated (Del Río González et al. 2011). Given their systemic nature, environmental-related innovations have to comply with multiple technical-economic problems, are expected to satisfy a variety of needs (Florida 1996; Oltra and Jean 2005), and more stringent regulatory requirements (Carrillo-Hermosilla et al. 2010). Similarly to other complex technologies, GTs, being at the technological frontier, result from the integration of different and heterogeneous technologies and knowledge sources (Petruzzelli et al. 2011). Their development is characterized by a substantial lack of standards in term of known accepted solutions, often requiring knowledge and skills distant from the traditional capabilities (De Marchi 2012; Fusillo et al. 2019). Moreover, recent empirical evidence suggests that GTs are often the results of the hybridization of green and dirty technologies in new and creative ways (Zeppini and Den Bergh 2011; Dechezleprêtre et al. 2015; Colombelli and Quatraro 2019) Because of the intrinsic complexity of environmental innovations, innovators may not possess internally the required competences and different knowledge resources to develop GTs. Hence, new and successful green innovations are more likely to emerge as an outcome of the collective efforts of organizations, research institutions, universities and teams of inventors. Several works, indeed, show that creating a wide net of collaborations with external partners is increasingly becoming essential for successful innovative performances, allowing to access and share specialized knowledge components residing outside traditional domains (De Marchi 2012; Fusillo et al. 2019; Petruzzelli et al. 2011). In a study on Spanish firms, De Marchi (2012) found that firms active in green innovative activities tend to cooperate with external partners to a greater extent than other innovative firms. Cainelli et al. (2015), confirming this result, suggest that internal, external and hybrid resource all are of greater importance in the introduction of new green product and processes. Orsatti et al. (2020) show that, on average, successful green inventions are more likely to be generated by team-level work, creatively recombining previous knowledge in original and unprecedented ways. In addition, recent contributions show that interactions with university and research institutes play a key role in the development of green technologies (Cainelli et al. 2012; De Marchi and Grandinetti 2013). Quatraro and Scandura (2019) found that the involvement of academic inventors exerts a positive impact on the generation of GTs because their strong scientific background enables to manage recombination across different technological domains.

Despite this abundance of studies, the empirical literature explicitly addressing the technological peculiarities of green recombination and extent to which they differ from more "traditional" technologies is relatively scant. At the technology level, Dechezleprêtre et al. (2015) compared the intensity of knowledge spillovers in clean and dirty technologies in energy production and transportation, showing that green patents receive more citations than their dirty counterparts and tend to be relatively more general and original. Similarly, Barbieri et al. (2020) focusing on both the search and the impact space in green and non-green technologies confirm that GTs tend to be more complex and radical than non-green ones and exert a higher impact on subsequent inventions.

The present study carries out a finer-grained comparison between the characteristics of green and non-green technologies. The aim is to test for the intrinsic complexity of green recombinations at the invention-technology level, focusing on two distinct – and yet interrelated – phases of the inventive process, i.e. the knowledge search phase and the knowledge production phase. We expect that the sources of knowledge from which GTs emerge are more likely to span technological fields that do not necessarily share significant commonalities, thus resulting in higher technological diversity during the search phase. Similarly, the diversified search, the hybridization of cognitively distant knowledge pieces, and their successful and creative integration are likely to lead to a higher degree of diversity of the effectively recombined knowledge. Therefore, the GTs invention process is characterized by a higher diversification in the knowledge production phase. According to the proposed arguments, we formulate the following working hypotheses:

H1: *Compared to non-green inventions, GTs exhibit higher technological diversity of the knowledge sources.*

H2: *Compared to non-green inventions, GTs exhibit higher technological diversity of the recombined knowledge.*

2.2 Technological diversity

According to the recombinant framework discussed above new technologies arise from the combination – and integration – of the multiple "bits" of knowledge cumulating in the knowledge bases. Scholars argued that the size and the interdependence of the knowledge base greatly affect the likelihood of a successful search. Also, the kind of performed search processes influences the risk of failures and the direction of technological development. New types of combination and new and distant components, increasing the variability and the uncertainty of the invention process, may lead to a higher rate of failure but also to breakthrough and radical invention (Fleming 2001; Nooteboom 2000)

This depiction of the knowledge generation process led researchers to stress the existence of a trade-off between the concepts of diversity and similarity. On the one hand, similarity in the knowledge bases eases communication and technological learning, facilitating the acquisition and integration of heterogeneous resources (Cantner and Meder 2007). On the other hand, a too strong overlap of competences may hamper the inventive process, restricting the scope of potential recombination, that may eventually lead to cognitive lock-in (Nooteboom 2000).

In empirical studies, the concept of technological diversity/similarity has been operationalized in different ways and by complementary measures. One of the most popular diversity measures is the Simpson index – usually known in the economics literature as Herfindahl-Hirschman concentration index – which measure the degree of concentration when the given elements are classified into types. Recently, a widely used index is represented by the technological variety, often decomposed into its two constituting components, the related variety and the unrelated variety. Computed by using the information entropy index (Shannon 1948), variety measures the extent of diversification in the knowledge base, within narrow technological areas – Related Variety (RU) – and across technologies – Unrelated Variety (UV) – (Frenken et al. 2007).[1]

Notable advancements in the field of Science and Interdisciplinarity studies provide a more comprehensive view of diversity, highlighting the importance of considering the intrinsic difference between knowledge components in constructing diversity indexes. Following the framework proposed by Rafols and Meyer (2010), diversity is a conceptual construct that "describes the difference in the bodies of knowledge that are integrated" (Rafols 2014). According to this framework, diversity can be characterized by three distinct attributes: i) the number of distinct categories into which element can be classified, i.e. *variety*; ii) the evenness of the distribution of the elements across the categories, i.e. *balance*; iii) the degree of difference between the categories, i.e. *disparity.* It turns out that an increase in diversity can be determined, independently, by an increase in each one of its attributes. To clarify, in the case of a patented invention, the diversity of such an invention increases along with the number of distinct technological classes to which it is assigned to, a more balanced distribution of those classes, and a higher difference (or technological distance) between those technologies. Although the decomposition into related and unrelated allows the variety index to partly take into account disparity by defining two sets with different disparity (Krafft et al. 2011), the entropy-based index better account for the number of distinct categories and the evenness of the distribution (Stirling 2007). Thus, in order to account also for the disparity attribute, Stirling (2007) proposed an integrated measure of diversity that weights the distribution of classified elements across their categories by their cognitive distance. First proposed by Rao (1982), the Rao-Stirling diversity index (or Integration Score) is increasingly used in empirical studies on knowledge integration and interdisciplinarity (Rafols 2014).

In order to stress the importance of weighting the distribution of technological classes by their cognitive distance – particularly in analyses at the invention-technology level –, we provide as an illustrative example two patent applications. The first application (patent number EP2551856B1), filed at the EPO in the 2011 by the Fujikura Ltd., presents the invention of a "high frequency cable and a high frequency coil which can suppress alternating-current resistance and can suppress heat generation and power consumption". The invention protected by the second patent application (patent number EP2598400B1) is filed in the same year at the EPO by the Davidson Tech Ltd., and relates to a "high altitude platform for gen-

[1] The entropy-based technological variety index, related and unrelated, is extensively employed in the Economic Geography literature to characterize sectoral and/or regional knowledge structure (see among others Boschma et al. (2012), Content and Frenken (2016), and Quatraro (2010))

erating electrical energy and delivering information services at altitude, including telecommunications, observation and positioning services".[2] Exploring the technological content of the two patents, we notice that according to the CPC technological classification both patents are assigned to 8 different main technology classes (at the 4-digits level). In terms of diversity attributes, this implies they share the same degree of *variety*, that equals 8. For what concerns the *balance* attribute, the entropy index reveals that their diversity is at very similar levels (entropy-based diversity equals 2.7329 in the"high frequency cable" patent and 2.8553 in the "high altitude platform"). Both patents recombine quite unrelated technologies as indicated by the higher Unrelated Variety score with respect to that of Related Variety (the UV is about 2.5 in both patents, while the RV is at around 0.2). Thus, accounting only for *variety* and *balance*, one may conclude that the diversity of the recombined knowledge in the two inventions does not significantly differ. However, when considering more explicitly *disparity* in measuring technological diversity a slightly different picture emerges, at the patent level. The Rao-Stirling diversity of the first patent is 0.58, while for the second one is at 0.79, placing them, respectively, at the 3rd and the 1st top quintile of the diversity distribution in our sample of EPO patent applications from 1980 to 2012.[3] Therefore, according to the Integration Score, the "high altitude platform" patent is considerably more diversified than the first one. An explanation for this higher diversity score lies in the finer inspection of the technologies to which the patent is assigned. Indeed, the "high altitude platform" invention recombines very different – and cognitively more distant – technologies: it embraces technologies related to captive balloons, tethered aircraft and special materials for ropes and cables, to fuel cells for renewable energy, through gas-turbine plants, rotary generators and reactant storage and supply.

In view of this, in this paper, we measure the technological diversity of green and non-green inventions by means of the Rao-Stirling diversity index. To the best of our knowledge, this work represents the first attempt to apply the Integration Score to the technology field. Accounting simultaneously for the three discussed attributes of diversity and, indirectly, for the (intrinsic) distance between technologies, the Rao-Stirling diversity index allows us to give a consistent characterization of the recombination patterns behind green technological development, providing an accurate tool to coherently compare green with non-green inventions.

[2]The full bibliographic data content of the two patents can be accessed by using their reported patent number through the Espacenet patent search platform, available at `https://worldwide.espacenet.com/`. We report in Appendix A the front page of the two patent documents.

[3]Detailed information on our sample and on the construction of the Rao-Stirling diversity index are presented in Section 3.

3 Empirical framework

3.1 Data

Our main source of data is represented by the information contained in patent documents at the European level. Although it is widely acknowledged that the use of patent data may have some drawbacks in empirical studies (Griliches 1990; Pavitt 1985), they still provide a "window on the knowledge economy" (Jaffe and Trajtenberg 2002), representing a major source of data for studying the development of technological knowledge (Strumsky et al. 2012).

Patent data are extracted from the "*OECD, REGPAT Database, March 2018*" and the "*OECD, Citation Database, March 2018*", which collect all the patent applications filed to the EPO and under the PCT (Patent Co-operation Treaty) from 1977, and all the citation present in the EPO and PCT patent documents published from 1978 onward.[4] Each patent is associated with at least one (usually more) technological class indicating the subject to which the invention relates. In this work we exploit the technological subclass of the Cooperative Patent Classification (CPC)[5] at the 4-digit level (CPC-4).

Our analysis focuses on the specific subset of patents applications that belongs to the green domain. We classify patents as environmentally-related according to the OECD ENV-TECH (Haščič and Migotto 2015) classification, based on the International Patent Classification (IPC) and Collaborative Patent Classification (CPC). The search strategy identified 203.388 green patent applications out of a total of 3.109.044 patent applications. Figure 1 plots the dynamics of green patent applications and the total number of applications over our sampled period 1980-2012. The figure shows that both the total number of applications and green inventions steadily increased over time at similar rates. Interestingly, from the 2000s onward we may notice a deep acceleration in the development of green technologies, which increased at a much faster pace.

The identified set of environmental-related patents includes quite heterogeneous inventions, addressing environmental challenges from different perspectives. According to the OECD ENV-TECH classification (Haščič and Migotto 2015), three macro environmental technology domains can be identified: i) Climate change mitigation technologies (related to energy, greenhouse gases, transport and building); ii)

[4]To account for incomplete counting and usual delays in the filing process we restricted the time-span to 1980-2012 for patent applications and to 1985-2012 when dealing with patent citations.

[5]The CPC is a new patent classification system, jointly developed by the European Patent Office (EPO) and United States Patents and Trademark Office (USPTO). Based on the European classification system (ECLA), it is a more detailed version of the International Patent Classification (IPC).

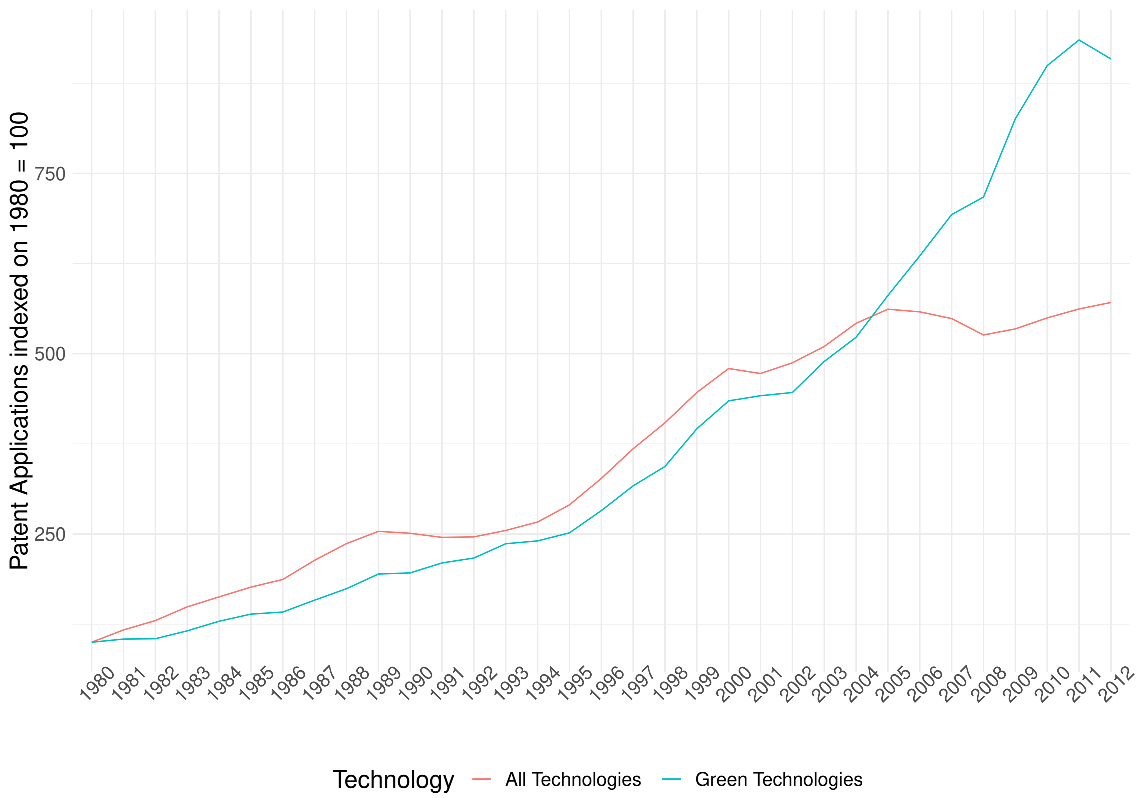

Figure 1: Dynamics of patent applications in all technologies and in Green Technologies. Note: Indexed on base year 1980 = 100

Environmental Management technologies (related to air and water pollution, waste management and soil remediation); iii) Water-related adaptation technologies (related to water scarcity). Therefore, we disentangle the dynamics shown in Figure 1 by decomposing the series into the three macro domains in which green inventions can be classified. As we can see from Figure 2 the fastest-growing technologies are those related to Climate Change Mitigation, which, among others, include renewable energy (wind, solar panels, etc.), and electric and hybrid vehicle technologies. On the other hand, technologies related to Environmental Management and Water-related Adaptation have been growing at a slower rate, more or less as much as inventions overall.

3.2 Measuring diversity

3.2.1 Rao-Stirling index

As anticipated in Section 2.2, we measure diversity using an indicator that allows us to account for the three distinct attributes of diversity, i.e. variety, balance and disparity. In its simplest formulation, the Rao-Stirling diversity index is given by:

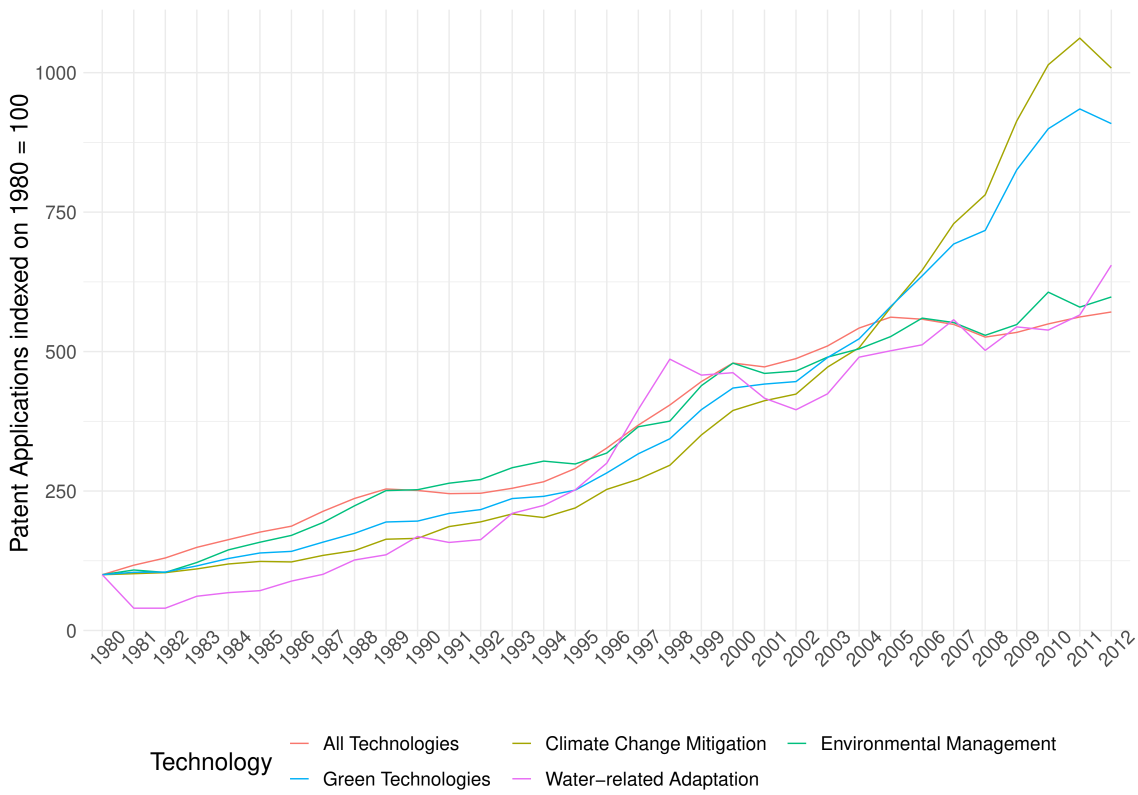

Figure 2: Dynamics of patents applications in selected environmental-related areas. Note: Indexed on base year 1980 = 100

$$\Delta_a = \sum_{i,j(i \neq j)} p_i p_j d_{ij} \tag{1}$$

where p_i and p_j are the proportions of element i and element j, respectively, and d_{ij} is the cognitive distance between the two. Thus, the Rao-Stirling can be interpreted as the average cognitive distance between the system's elements. Since our unit of analysis is the patent, we proxy system elements with the different technological classes to which the patent is assigned (co-classification). In turn, the technological diversity of patent a, is given by the relative proportion of the 4-digit CPC classes within the patent, weighted by their cognitive distance. The co-citations of patent classes are instead used to compute of the diversity of the knowledge sources, meaning that the diversity depends on the relative proportion of technological classes within the backward citations of the focal patents.

3.2.2 Cognitive distance and normalization

While computing the distribution of classes is straightforward, the choice of the appropriate cognitive distance measure between technologies can be challenging. Technological or cognitive distance (often known in existing literature as proximity or relatedness) can be measured in several ways, each of them reflecting slightly different intuitions on how technologies are related. Considering only measures computed using patent data, we can distinguish between two main broad "families": citations and co-classification measure.[6] Given that backward citations reveal prior art, a first straightforward measure of proximity is given by the count of citations from patents in a class i to patents in class j (Leten et al. 2007). Another common measure is based on the concept that if two classes are often cited together it may imply that they possess some degree of relatedness. Referred to as co-citation, it counts the number of patents that cited patents classified in both class i and j (Uzzi et al. 2013; Wallace et al. 2009). Elaborating upon the co-citation, another well-known measure is based on cosine similarity between the vectors counting how many citations patents in each class make to every other class. Concerning the co-classification family, the most straightforward proximity measure consists in calculating the frequency with which two classes are assigned to a patent (Dolfsma and Leydesdorff 2011; Engelsman and Raan 1994; Jeong et al. 2015; Joo and Kim 2010; Leydesdorff 2008). Similarly, co-occurrence based technological proximity can be derived by measuring how often two classes appear together in the patenting histories of "actors" – e.g. inventors, firms, region, countries etc. Belongs to this family the relatedness measure first introduced by Jaffe (1986), computed using the cosine similarity between the occurrence of classes in firm's patents. Then, this approach has been extended and used for general vectors of co-occurrence of technological class (Breschi et al. 2003; Kogler et al. 2013). More recently, following the contribution of Hidalgo et al. (2007), a new group of measures, based on the diversification behaviour of actors, emerged in the literature. These indexes measure the proximity between two technological classes in terms of the likelihood of an actor to develop a Revealed Technological Advantage (RTA) in one class, given that he has already developed an RTA in the other.

A recent stream of technical literature dealing with technological proximity claims that all these measures tend to be affected by factors not intrinsically related to the technologies themselves, that may, in turn, distort the true information (Alstott et al. 2017). For example, the likelihood of two class co-occurring in the same patents depends both on the total number of classes in which a patent is

[6] For an extensive review on the various patent-based distance measures see Yan and Luo (2017).

classified and on the number of patents associated with a given technology. Teece et al. (1994) introduced a normalization method (null model) in order to account for the class size effect, by computing the deviation of the observed co-occurrences from their expected values.[7] However, Bottazzi and Pirino (2011) suggest that in order to obtain a measure of "true" proximities empirical co-occurrence should be compared with a null hypothesis in which are preserved both the number of classes per patent and the number of patents per class. Alstott et al. (2017) expanded this normalization by controlling also for the temporal effect, as occurrences may significantly vary over time.

In the present work, we measure technological proximity by using the normalized co-occurrences between 4-digits CPC codes. To do so, we proceed in a number of steps. First, from the occurrence of classes in each patent, we construct a symmetric co-occurrence matrix C in which each cell C_{ij} represents the number of patents associated with both technology i and j. At this point, to obtain an unbiased proximity measure between technology, we compare the empirical co-occurrence matrix with randomized controls generated by a null model as discussed in Alstott et al. (2017). We first create 1000 randomized versions of the patent-class occurrences for each year from 1980 to 2012, preserving both the number of classes per patent and the number of patents per class. Then, we combine the yearly versions into a single randomized version of the co-occurrence matrix. By dividing the empirical co-occurrence matrix C by the randomized one we obtain a new matrix E in which each cell E_{ij} represents the deviation of the observed number of patents in both technology i and j from its expected value. Lastly, the normalized proximity measure P between all technologies is derived by applying the *cosine index* to the normalized co-occurrence matrix E, in order to obtain a measure ranging between 0 and 1. The cognitive distance used to weight the proportion of combined technologies in our Rao-Stirling diversity index is simply obtained by:

$$d_{ij} = 1 - p_{ij} \tag{2}$$

Measures of proximity have also been widely used to develop global maps of the knowledge space which can provide valuable information about promising areas and

[7]In Teece et al. (1994) the expected value is calculated by randomly assigning firms to industrial sectors.

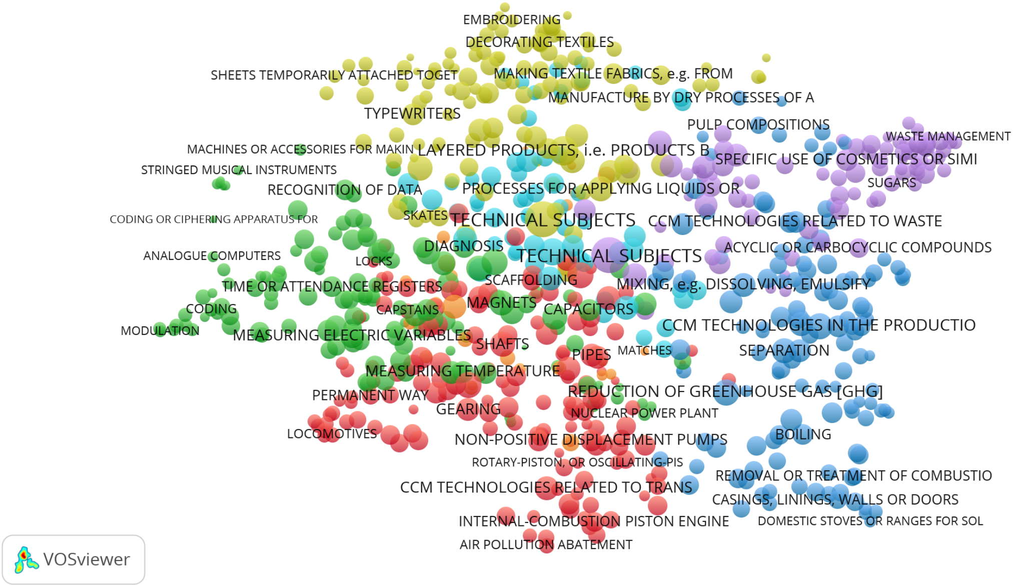

Figure 3: Map of the technology space according to the normalized measure of proximity, at CPC 4-digit level. Nodes are colored according to the community structure. To ease visualization, links have been hidden.

the positioning of specific entities.[8] As an example, we plot in Figure 3 the map of the technology space derived from our normalized technological proximity measure. The map of technological space is graphically represented as a network in which each node is a technological class and links represents their proximity.

3.3 Methodology

3.3.1 Propensity score matching

Our analysis aim at comparing the degree of diversity of green *vis-à-vis* non-green inventions. This task requires the identification of a coherent "control group" consisting of non-green inventions which are expected to be similar to the green ones. We identify the control group by performing a propensity score matching, as it allows us to take into account all the technological areas covered by green inventions. The propensity score measures the probability of a patent belonging to the green technology domain, given a set of patent-level observable characteristics. It is recovered through the estimation of a probabilistic choice model where the dependent variable equals 1 when the patent is green and 0 otherwise. Given the binary nature of the dependent variable, to estimate the probabilistic model we implement a probit model. For what concerns the choice of the covariates, we follow prior literature (Appio

[8]See, for example, the global map of science in Leydesdorff and Rafols (2009).

et al. 2017) and we include in the model a rich set of predictors at the patent level: the *patent scope*, defined as the number of distinct technology classes the invention is assigned to; the *family size*, proxied by the number of patent offices at which a given invention has been protected; the number of *backward citations*; the number of citations to *non-patent literature* (e.g. scientific paper, conference proceeding, databases); the number of *claims* included in the patent; the number of citations received (*forward citations*); the number of *inventors*; the number of *applicants*. We also include in the model a set of dummies controlling for the macro geographical region from which the patents originate, exploiting the information on the inventors' addresses.[9] Lastly, we split the raw sample of patents into 35 sub-samples on the basis of the priority year assigned to each patent application and performed the propensity score matching individually for each year. The 1:1 matching allowed us to select a control group of 203.388 non-green patents (out of 2.905.656 potential controls) matched with the same number of green patents. Details on the matching outcome and covariates balance are presented in Table 1, which describes the variables employed to estimate the propensity score and the mean comparison before and after the matching. Before the matching patent characteristics differ substantially between the green and non-green patents, supporting the need for selecting an ad-hoc group. However, the propensity score matching yields satisfactory results, as the difference in means between the two groups significantly decreases after the matching.[10]

3.3.2 Green VS non-green

To empirically test the difference between the green and the control group of non-green inventions we estimate the following simple model:

$$\begin{aligned}\Delta_i = \alpha + \beta_1 \text{greenPat}_i + \beta_2 \text{inventors}_i + \beta_3 \text{applicants}_i \\ + \beta_4 \text{backCitations}_i + + \beta_5 \text{familySize}_i + \text{IPC35} + \omega_i + t_i + \varepsilon_i\end{aligned} \tag{3}$$

where Δ_i, the dependent variable, is the diversity of patent i or the diversity in its patent citations. greenPat_i is our focal explanatory variable, and it is a dummy

[9]We identified 16 macro-regions: Australia and New Zealand, Central Asia, Eastern Asia, Southern Asia, South-eastern Asia, Western Asia, Eastern Europe, Northern Europe, Southern Europe, Western Europe, Northern America, Latin America and the Caribbean, Micronesia, Polynesia, Northern Africa, Sub-Saharan Africa.

[10]An identical propensity score matching procedure is performed to select the control group for the analysis of diversity in patent citations. The procedure allows to match 103.161 green patents with the same number of non-green ones. This reduction in the sample size is due to the exclusion of patents for which it has not been possible to retrieve the technology classes assigned to its cited patents. The comparison between pre-matching and post-matching yields similar satisfactory results.

Table 1: Mean comparison of patent characteristics between Green patents (N = 203.388) and controls patents (raw sample 2.905.656) before and after the Propensity Score Matching.

	Mean before Matching		**Mean post Matching**	
	Controls	Green	Controls	Green
Patent Scope	1,886695	2,130042	2,191157	2,184313
Family Size	5,527260	5,413714	5,531344	5,641542
Backward Citations	5,502565	6,179672	5,894114	5,961517
Non-patent Literature	1,523676	1,617449	1,512002	1,531093
Claims	13,219684	12,840990	12,697126	12,903529
Forward Citations	1,082778	1,199253	1,325526	1,430345
Inventors	2,579484	2,694697	2,595871	2,660108
Applicants	1,077386	1,091549	1,084843	1,090831

variable taking value 1 if the patent is classified as green and 0 otherwise. The model includes controls for the number of distinct inventors and applicants in each patent. backCitations_i and familySize$_i$, control, respectively, for the total number of citations made by the focal patent and the size of the patent family, as factors that may be associated to the degree of diversity. To account for potential geographical heterogeneity and time-varying effects, we include a set of dummies controlling for 16 world macro-regions (ω_i) and a set of time dummies corresponding to a 5-year time window (t_i). To rule out any possible heterogeneity in technology areas remaining after the propensity score matching, we included a control for the main technology domain assigned to each patent based on the 35-technology fields as in Schmoch (2008) (IPC35). Our dependent variable, the Rao-Stirling technological diversity, is bounded between 0 and 1 by construction. Existing empirical literature does not provide homogeneous consensus regarding the best econometric specification with such dependent variables (Argyres and Silverman 2004; Laursen and Salter 2006). Hence, the main estimations are carried out through OLS regressions to ease interpretation. A robustness check employing censored Tobit regressions as an alternative model is reported in Section 4.3.

4 Results

Descriptive statistics of the variables of interest are reported in Table 2. Figure 4 and Figure 5 show the evolution of the average diversity of green and non-green inventions, respectively in the production phase and the search phase. From a visual

Table 2: Descriptive Statistics

	N	Mean	St. Dev.	Min	Max
Diversity	406,776	0.29612	0.20733	0	0.82074
Inventors	406,776	2.62799	1.88469	1	60
Applicants	406,776	1.08784	0.38829	1	18
BackCitations	406,776	5.92782	9.32448	0	1,011
FamilySize	406,776	5.58644	3.93422	1	55
Claims	406,776	12.80033	9.30636	1	442

inspection, it is evident that the average diversity in green patents is significantly and persistently higher than the diversity in the control group. Diversity in green recombinations slightly decreases during the first 15 years of our sample, while a clear increasing trend is identifiable from 2000 onward. On the other hand, the diversity in the knowledge sources of green invention increased through all the period considered. Figure 4 and 5 provide preliminary interesting insights on the more diversified nature of knowledge that characterize the recombination activities of green inventions.

4.1 Are GTs more diversified?

Table 3 reports the main results of our empirical analysis. Columns 1-3 test whether green inventions do exhibit higher levels of technological diversity in the recombined knowledge, thus focusing on the knowledge production phase. Columns 4-6 refer to the diversity in the recombined knowledge sources, thus estimating the diversity premium of green inventions during the knowledge search phase. Controls are gradually added in models (2) and (3), and models (5) and (6). All models include time, regional and main technology effects, and reported standard errors are heteroskedastic-robust.

For what concerns the diversity in the production phase, the green patent dummy has a positive and strongly significant coefficient across all specifications. The positive and significant coefficients on *inventors* and *applicants* in both columns 2 and 3 suggest that, on average, teams of 2 or more inventors (and multiple applicants) are able to generate inventions with higher technological diversity. This is in line with existing literature showing that collaborations allow to exploit synergies and knowledge exchanges between partners, resulting in highly diversified and more valuable inventions. The GTs premium in terms of technological diversity is quite stable at around 0.20 also when controlling for the number of backward citations, the number of claims made in the patent and the family size, which all are significantly associated with higher diversity in recombination. This first set of results confirms

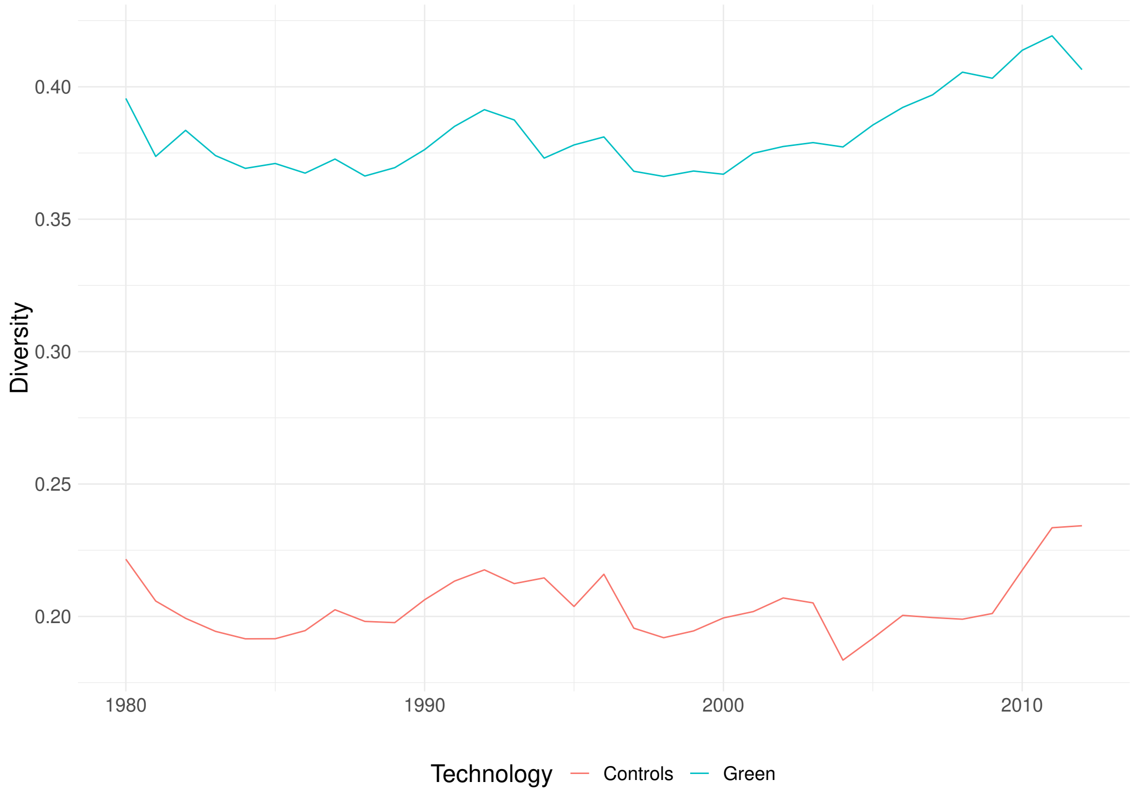

Figure 4: Average diversity of recombination (knowledge production phase) in green and non-green inventions (control group)

our first hypothesis, indicating that, on average, GTs do recombine more diversified pieces of knowledge as compared to similar but non-green inventions.

Consistent with H2 results reported in columns 4-6 of Table 3 indicate that, in comparison to non-green inventions, GTs draw from more diversified knowledge sources, thus exhibiting higher technological diversity in the knowledge search phase. The green premium is positive and statistically significant, with an estimated coefficient slightly lower in magnitude. The result holds when controls are added to the model. Controls retain their sign and significance, with the exclusion of the number of applicants and patent family size, for which no statistically significant coefficients are found.

4.2 Are all GTs domains more diversified?

The results presented in Section 4 confirm that green inventions, on average, both recombine more diversified technologies and their knowledge sources span a wide range of distant technology areas. However, as anticipated in Section 3.1 the set of GTs itself includes technologically heterogeneous inventions that try to address environmental challenges from different perspectives. Consequently, in order to test

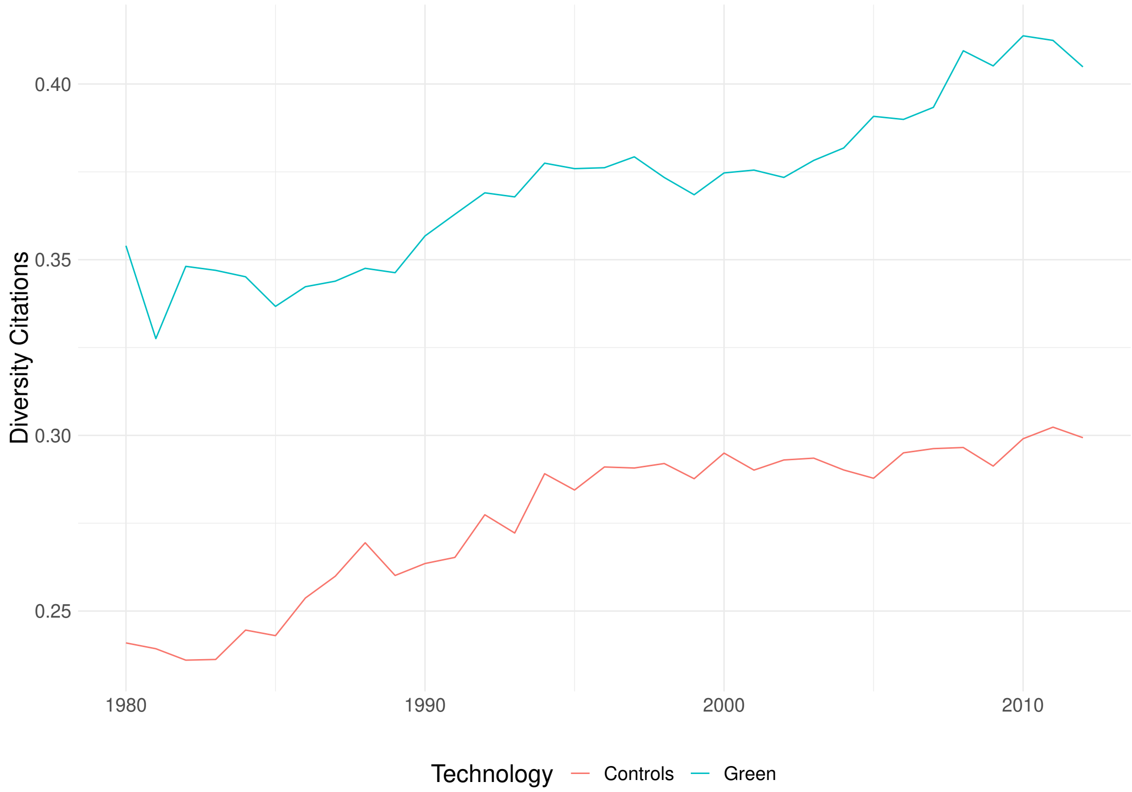

Figure 5: Average diversity of knowledge sources (knowledge search phase) in green and non-green inventions (control group)

whether and to what extent the green premium may differ among environmental domains, we categorize green inventions as belonging to the three macro environmental categories and split the sample of patents accordingly. Following the same procedure employed for the whole sample, we identified an ad-hoc control group on non-green inventions for each set of patents. Then, we re-run our regression analysis in order to estimate whether technological diversity in recombination differs between green and non-green inventions, within the three macro environmental domains.

Results of econometric estimation for the three selected environmental domains are reported in Table 4. Columns 1-3 shows that the diversity premium of Climate change mitigation technologies reaches 0.22 and it is statistically significant across the specifications. Environmental Management technologies, also, are significantly associated with a higher degree of diversity compared to their control group, with slightly lower a coefficient (columns 4-6). Lastly, the premium in terms of diversity of Water-related adaptation technologies reduces at around 0.12 (columns 7-9), still with statistically significant coefficients. In sum, the findings of Table 4 suggest that, when categorizing green inventions in macro environmental domains, GTs still exhibit a higher degree of diversity in the technological recombination process, though at different magnitudes. Interestingly, the higher premium found for Climate change mitigation technologies over Environmental Management and Water-related adaptation – and Environmental Management over Water-related adaptation – is consistent with the dynamics of green patent application shown in Figure 2. Indeed, climate change mitigation technologies, by including technologies related to, among other, renewable energy or electric vehicle, experienced the fastest rate of development, particularly from the 2000s onward.

Table 3: OLS Regression Results of GTs on Technological Diversity with geographical, time and main technology controls.

	Diversity in recombination			Diversity in knowledge sources		
	(1)	(2)	(3)	(4)	(5)	(6)
GreenPat	0.20032***	0.20013***	0.19853***	0.08530***	0.08501***	0.08384***
	(0.00068)	(0.00068)	(0.00068)	(0.00107)	(0.00107)	(0.00107)
Inventors		0.00176***	0.00103***		0.00186***	0.00112***
		(0.00016)	(0.00016)		(0.00024)	(0.00025)
Applicants		0.00196***	0.00226***		0.00197	0.00181
		(0.00074)	(0.00074)		(0.00121)	(0.00120)
BackCitations			0.00041***			0.00174***
			(0.00003)			(0.00004)
FamilySize			0.00193***			−0.00018
			(0.00008)			(0.00013)
Claims			0.00068***			0.00069***
			(0.00003)			(0.00005)
Constant	0.19829***	0.19351***	0.17088***	0.21729***	0.21198***	0.19686***
	(0.00340)	(0.00349)	(0.00354)	(0.00556)	(0.00571)	(0.00580)
Time Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region Controls	Yes	Yes	Yes	Yes	Yes	Yes
IPC35 Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	406,775	406,775	406,775	206,321	206,321	206,321
R^2	0.24027	0.24053	0.24298	0.09875	0.09903	0.10677
Adjusted R^2	0.24017	0.24042	0.24287	0.09851	0.09878	0.10651
F Statistic	2,338.66200***	2,259.80400***	2,175.68900***	410.91630***	397.74550***	410.93280***

Note: Heteroskedastic-Robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 4: OLS Regression Results of selected environmental domains on Technological Diversity with geographical, time and main technology control

	Diversity in recombination								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
ClimateChange	0.21937***	0.21927***	0.21841***						
	(0.00081)	(0.00081)	(0.00081)						
EnvManagement				0.18268***	0.18249***	0.17948***			
				(0.00155)	(0.00155)	(0.00155)			
WaterRelated							0.11678***	0.11708***	0.11807***
							(0.00346)	(0.00346)	(0.00345)
Inventors		0.00077***	0.00029		0.00283***	0.00197***		0.00323***	0.00225***
		(0.00019)	(0.00019)		(0.00031)	(0.00031)		(0.00056)	(0.00056)
Applicants		0.00021	0.00042		0.00188	0.00241*		−0.00145	−0.00070
		(0.00089)	(0.00089)		(0.00132)	(0.00131)		(0.00171)	(0.00170)
BackCitations			0.00032***			0.00129***			0.00008
			(0.00004)			(0.00010)			(0.00006)
FamilySize			0.00104***			0.00294***			0.00227***
			(0.00009)			(0.00016)			(0.00025)
Claims			0.00056***			0.00096***			0.00062***
			(0.00004)			(0.00007)			(0.00009)
Constant	0.20172***	0.20038***	0.18481***	0.19349***	0.18758***	0.14819***	0.19485***	0.19097***	0.17245***
	(0.00444)	(0.00454)	(0.00460)	(0.00659)	(0.00674)	(0.00687)	(0.01325)	(0.01338)	(0.01344)
Time Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IPC35 Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	243,929	243,929	243,929	138,611	138,611	138,611	24,234	24,234	24,234
R^2	0.30771	0.30776	0.30900	0.17259	0.17313	0.17776	0.13080	0.13202	0.13685
Adjusted R^2	0.30755	0.30759	0.30883	0.17226	0.17279	0.17740	0.12886	0.13001	0.13474
F Statistic	1,970.84300***	1,902.10200***	1,817.51900***	525.47240***	508.95280***	499.21720***	67.38172***	65.66423***	64.95971***

Note: Heteroskedastic-Robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

4.3 Robustness checks

In this section, we report some additional estimates in order to check for the robustness of our empirical analysis. First, since our dependent variable is, by construction, bounded between 0 and 1, we test for the robustness of our results to the choice of the estimation model, by re-estimating our main model using left-censored Tobit Regressions.[11] Tobit regressions estimates, presented in Table 5, largely confirm our previous results. Columns 1-3 refer to diversity in the knowledge production phase, while columns 4-6, refer to the diversity in knowledge sources. In line with the OLS estimations, technological diversity is significantly higher in green patents as compared to non-green and the diversity in recombination is higher in magnitude with respect to the diversity in citations.

In Section 2.2 we extensively stress the importance of measuring technological diversity by considering all its intrinsic attributes, especially when working at the narrow technology-invention level. Nevertheless, we expect that the direction of the hypothesized relationships should be, at least in principle, robust to the choice of the diversity measure, as long as 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 do so, we include a dependent variable in the regressions the Entropy index and its decomposition into Related Variety (RV) and Unrelated Variety (UV). The three alternative measures are normalized between 0 and 1 by dividing by $log_2(N)$ – where N is the total number of existing technology classes – in order to obtain coefficients comparable in magnitude to those of our main estimation. The results are presented in Table 6 and, overall confirm our main findings. As reported in columns 1-3, GTs are associated with higher levels of entropy-based diversity, in comparison to non-green inventions. Similar results hold when RV (columns 4-6) and UV (columns 7-9) are employed as alternative diversity measures. Interestingly, the green diversity premium in terms of UV is higher in magnitude compared to that in terms of RV, and it is very close to the estimated coefficient in the entropy-based diversity model, suggesting that the higher overall diversity of GTs seems to be mainly guided by unrelated diversification.

[11] It is worth stressing that, since the theoretical maximum of the dependent variable (equal to 1) is never reached in the data, therefore, the model is not right-censored.

Table 5: Tobit Regression Results of GTs on Technological Diversity with geographical, time and main technology controls.

	Diversity in recombination			Diversity in knowledge sources		
	(1)	(2)	(3)	(4)	(5)	(6)
GreenPat	0.23987***	0.23963***	0.23922***	0.11846***	0.11801***	0.11769***
	(0.00075)	(0.00075)	(0.00074)	(0.00108)	(0.00108)	(0.00108)
Inventors		0.00376***	0.00251***		0.00351***	0.00219***
		(0.00020)	(0.00020)		(0.00029)	(0.00029)
Applicants		0.00396***	0.00374***		0.00652***	0.00576***
		(0.00097)	(0.00094)		(0.00147)	(0.00146)
BackCitations			0.00073***			0.00224***
			(0.00003)			(0.00005)
FamilySize			0.00293***			−0.00009
			(0.00008)			(0.00015)
Claims			0.00092***			0.00066***
			(0.00003)			(0.00006)
Time Controls	Yes	Yes	Yes	Yes	Yes	Yes
Region Controls	Yes	Yes	Yes	Yes	Yes	Yes
IPC35 Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	406,775	406,775	406,775	206,321	206,321	206,321
Log-likelihood	-85371.68	-85185.58	-84279.59	-43546.18	-43457.08	-42462.82

Note: Heteroskedastic-Robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table 6: OLS Regression Results of GTs on Entropy, Related and Unrelated Variety, with geographical, time and main technology controls.

	Dependent variable:								
	Entropy			RV			UV		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GreenPat	0.07322***	0.07307***	0.07228***	0.00965***	0.00959***	0.00945***	0.06356***	0.06348***	0.06283***
	(0.00025)	(0.00025)	(0.00025)	(0.00015)	(0.00015)	(0.00015)	(0.00022)	(0.00022)	(0.00022)
Inventors		0.00132***	0.00096***		0.00062***	0.00055***		0.00070***	0.00041***
		(0.00006)	(0.00006)		(0.00003)	(0.00003)		(0.00005)	(0.00005)
Applicants		0.00110***	0.00125***		0.00008	0.00010		0.00102***	0.00115***
		(0.00027)	(0.00027)		(0.00016)	(0.00016)		(0.00024)	(0.00024)
BackCitations			0.00022***			0.00009***			0.00013***
			(0.00001)			(0.00001)			(0.00001)
FamilySize			0.00094***			0.00013***			0.00081***
			(0.00003)			(0.00002)			(0.00003)
Claims			0.00034***			0.00008***			0.00026***
			(0.00001)			(0.00001)			(0.00001)
Constant	0.06113***	0.05794***	0.04667***	0.00374***	0.00271***	0.00049	0.05739***	0.05523***	0.04618***
	(0.00126)	(0.00129)	(0.00131)	(0.00074)	(0.00076)	(0.00078)	(0.00112)	(0.00115)	(0.00117)
Time Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IPC35 Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	406,775	406,775	406,775	406,775	406,775	406,775	406,775	406,775	406,775
R^2	0.27416	0.27515	0.27942	0.09901	0.09974	0.10057	0.25468	0.25506	0.25851
Adjusted R^2	0.27407	0.27505	0.27932	0.09889	0.09961	0.10044	0.25457	0.25496	0.25840
F Statistic	2,793.21800***	2,708.51800***	2,628.59500***	812.65530***	790.53990***	757.97270***	2,526.81800***	2,443.14400***	2,363.30500***

Note: Heteroskedastic-Robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01

5 Conclusions

By exploiting a rich dataset of European patent applications from 1980 to 2012, the present paper investigates the recombination patterns of Green Technologies in order to assess whether and to what extent the generation of green knowledge involves higher recombinatorial complexity with respect to more "traditional" inventions. Based on prior evidence, we hypothesize that the development of GTs is more likely to require knowledge and skills distant from traditional knowledge bases, which, in our framework, translates into a higher technological diversity in comparison to non-green inventions. Two early phases of the invention process are taken into account: the knowledge search phase, where the diversity in technological content of source of knowledge is considered, and knowledge production phase, where the focus is on the pieces of knowledge effectively recombined. The paper operationalizes the concept of technological diversity by using a comprehensive measure of diversity, which weights the co-occurrence of technologies within patented inventions with a normalized measure of cognitive distance between the technologies. An ad-hoc set of non-green patents is identified through a 1:1 propensity score matching and employed as control group in the empirical analysis. We find evidence that green inventions are more likely to recombine more diverse pieces of knowledge, spanning a wider range of technology areas. Empirical results also corroborate our second hypothesis according to which, during the search phase, green technologies draw upon more diversified knowledge sources, involving technologies which are often cognitively far from each other. In addition, in both phases of the invention process, diversity is positively influenced by the number of inventors and applicants, confirming that the involvement inventors teams and pooled applicants, often with different backgrounds, may allow the recombination of different and distant technological domains, generating more complex and (possibly) successful inventions.

Our study is not free from caveats. The first limitation relates to the use of patent data and technological classes in properly identifying green technologies. Patents represent a relevant but partial subset of the wide range of the types and forms of knowledge that are relevant for scientific, technological and economic advance. However, the fine-grained information contained in patent documents still represent a highly reliable tool to study technological dynamics (Griliches 1981; Strumsky et al. 2012). A second limitation concerns the geographical dimension of green technological development. We use the information of inventors address to geo-localize patents and include controls for the broad regions the invention originates from. Yet, notwithstanding the increasing worldwide efforts in raising awareness of the environmental challenges, regions and countries may still highly differ in terms of policy response. This, in turn, affects their green technological advancements and the propensity toward greener innovative solution, as confirmed by the vast literature on the effect of stringent environmental policies on environmental innovations (Arimura et al. 2007; Carrión-Flores and Innes 2010; Ghisetti and Quatraro 2015; Johnstone et al. 2010; Lanoie et al. 2011; Popp 2003). Future research might adopt explicitly an environmental policy perspective and investigate the relationship between region/countries environmental policies and the technological complexity of green technologies. Third, our empirical framework does not allow to fully address potential concerns regarding the identification of a comparable set of non-green inventions. This is due, for example, to the possibility that, in addition to the ex-ante patent characteristics included in our propensity score matching, other factors, such as the development stage of technologies, might affect the correct identification of a control group. A comparison of inventions along their technological life cycle deserves further investigation going beyond the scope of this paper.

Nevertheless, this work importantly adds to the academic debate by moving a step further in the understanding of the antecedents of GTs and the dynamics underlying their generation. Building on recent academic developments on the intrinsic complexity of GTs, this paper digs into the technological content of the environmental invention. It provides novel empirical evidence on the higher technological diversity in green recombinations with respect to traditional inventions, confirming and supporting the importance of investigating GTs peculiarities.

Our findings lead to important policy implications. The work contributes to the debate on the optimal policy intervention to stimulate advancements in GTs in order to reduce environmental pressures and restore production efficiency and competitiveness. The capacity of green patents to collect different pieces of knowledge and to recombine them in a more effective way than other invention is a specific point to focus on when distributing public resources to foster innovation. The higher technological diversity found in green inventions supports the need to implement R&D policies targeted specifically at boosting environmental technologies. Moreover, the wide array of different and distant technologies recombined in GTs, calls for efforts in stimulating interactions and collaborations in order to assess external specialized knowledge. In line with recent policy instruments aimed at promoting interdisciplinarity in science, policymakers may find ways to further stimulate the generation of GTs by boosting the creation of network dynamics between heterogeneous actors, endowing them with the different competencies and skills required to successfully recombined diversified technological knowledge. Finally, a deeper identification of the more diversified technological domains from which the recombined knowledge in GTs comes from might also support the design of appropriate policy instruments to better exploit potential knowledge hybridizations and technology spillover effects.

References

Alstott, J., G. Triulzi, B. Yan, and J. Luo (2017). “Mapping technology space by normalizing patent networks”. In: *Scientometrics* 110.1, pp. 443–479.

Appio, F. P., A. Martini, and G. Fantoni (2017). “The light and shade of knowledge recombination: Insights from a general-purpose technology”. In: *Technological Forecasting and Social Change* 125, pp. 154–165.

Archibugi, D. and M. Pianta (1994). “Aggregate convergence and sectoral specialization in innovation”. In: *Journal of Evolutionary Economics* 4.1, pp. 17–33.

Argyres, N. S. and B. S. Silverman (2004). “R&D, organization structure, and the development of corporate technological knowledge”. In: *Strategic Management Journal* 25.89, pp. 929–958.

Arimura, T. H., A. Hibiki, and N. Johnstone (2007). “An Empirical Study of Environmental R&D: What Encourages Facilities to be Environmentally Innovative?” In: *Environmental policy and corporate behaviour*, pp. 142–173.

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

Boschma, R., A. Minondo, and M. Navarro (2012). “Related variety and regional growth in Spain”. In: *Papers in Regional Science* 91.2, pp. 241–256.

Bottazzi, G. and D. Pirino (2011). “Measuring Industry Relatedness and Corporate Coherence”.

Breschi, S., F. Lissoni, and F. Malerba (2003). “Knowledge-relatedness in firm technological diversification”. In: *Research Policy* 32.1, pp. 69–87.

Brunnermeier, S. B. and M. A. Cohen (2003). “Determinants of environmental innovation in US manufacturing industries”. In: *Journal of Environmental Economics and Management* 45.2, pp. 278–293.

Cainelli, G., V. De Marchi, and R. Grandinetti (2015). “Does the development of environmental innovation require different resources? Evidence from Spanish manufacturing firms”. In: *Journal of Cleaner Production* 94, pp. 211–220.

Cainelli, G., M. Mazzanti, and S. Montresor (2012). “Environmental Innovations, Local Networks and Internationalization”. In: *Industry and Innovation* 19.8, pp. 697–734.

Cantner, U. and A. Meder (2007). “Technological proximity and the choice of cooperation partner”. In: *Journal of Economic Interaction and Coordination* 2.1, pp. 45–65.

Carrillo-Hermosilla, J., P. Del Río, and T. Könnölä (2010). “Diversity of eco-innovations: Reflections from selected case studies”. In: *Journal of Cleaner Production* 18.10-11, pp. 1073–1083.

Carrión-Flores, C. E. and R. Innes (2010). “Environmental innovation and environmental performance”. In: *Journal of Environmental Economics and Management* 59.1, pp. 27–42.

Carrión-Flores, C. E., R. Innes, and A. G. Sam (2013). “Do voluntary pollution reduction programs (VPRs) spur or deter environmental innovation? Evidence from 33/50”. In: *Journal of Environmental Economics and Management* 66.3, pp. 444–459.

Colombelli, A. and F. Quatraro (2019). “Green start-ups and local knowledge spillovers from clean and dirty technologies”. In: *Small Business Economics*.

Content, J. and K. Frenken (2016). “Related variety and economic development: a literature review”. In: *European Planning Studies* 24.12, pp. 2097–2112.

Corrocher, N., F. Malerba, and F. Montobbio (2007). “Schumpeterian patterns of innovative activity in the ICT field”. In: *Research Policy* 36.3, pp. 418–432.

De Marchi, V. (2012). “Environmental innovation and R&D cooperation: Empirical evidence from Spanish manufacturing firms”. In: *Research Policy* 41.3, pp. 614–623.

De Marchi, V. and R. Grandinetti (2013). “Knowledge strategies for environmental innovations: The case of Italian manufacturing firms”. In: *Journal of Knowledge Management* 17.4, pp. 569–582.

Dechezleprêtre, A., R. Martin, and M. Mohnen (2015). “Knowledge spillovers from clean and dirty technologies : a patent citation analysis Centre for Climate Change Economics and Policy”.

Del Río González, P. (2009). “The empirical analysis of the determinants for environmental technological change: A research agenda”. In: *Ecological Economics* 68.3, pp. 861–878.

Del Río González, P., M. Á. Tarancón Morán, and F. C. Albiñana (2011). “Analysing the determinants of environmental technology investments. A panel-data study of Spanish industrial sectors”. In: *Journal of Cleaner Production* 19.11, pp. 1170–1179.

Dolfsma, W. and L. Leydesdorff (2011). “Innovation systems as patent networks: The Netherlands, India and Nanotech”. In: *Innovation: Management, Policy and Practice* 13.3.

Engelsman, E. C. and A. F. van Raan (1994). "A patent-based cartography of technology". In: *Research Policy* 23.1, pp. 1–26.

Fleming, L. (2001). "Recombinant uncertainty in technological search". In: *Management science* 47.1, pp. 117–132.

Fleming, L. and O. Sorenson (2001). "Technology as a complex adaptive system: Evidence from patent data". In: *Research Policy* 30.7, pp. 1019–1039.

Florida, R. (1996). "Lean and Green: The Move to Environmentally Conscious Manufacturing". In: *California Management Review* 39.1, pp. 80–105.

Frenken, K., F. Van Oort, and T. Verburg (2007). "Related variety, unrelated variety and regional economic growth". In: *Regional Studies* 41.5, pp. 685–697.

Fusillo, F., F. Quatraro, and S. Usai (2019). "Going Green: Environmental Regulation, eco-innovation and technological alliances". In: *Working paper series CENTRO RICERCHE ECONOMICHE NORD SUD (CRENoS)* 07, pp. 1–32.

Ghisetti, C., A. Marzucchi, and S. Montresor (2015). "The open eco-innovation mode. An empirical investigation of eleven European countries". In: *Research Policy*. Vol. 44. 5, pp. 1080–1093. ISBN: 0048-7333.

Ghisetti, C. and F. Quatraro (2015). "Regulatory push-pull effects on innovation : an evaluation of the effects of the REACH regulation on patents in the chemical sector". In: *WWWforEurope Working Paper no 91* 91.

Griliches, Z. (1981). "Market value, R&D, and patents". In: *Economics Letters*.

— (1990). *Patent Statistics as Economic Indicators: A Survey*. ISBN: 00220515.

Haščič, I. and M. Migotto (2015). "Measuring environmental innovation using patent data". In: *OECD Environment Working Papers* No. 89.89, p. 59.

Hidalgo, C. A., B. Winger, A. L. Barabási, and R. Hausmann (2007). "The product space conditions the development of nations". In: *Science* 317.5837, pp. 482–487.

Jaffe, A. B. (1986). "Technological Opportunity and Spillovers of R & D: Evidence from Firms' Patents, Profits, and Market Value". In: *The American Economic Review*.

Jaffe, A. B. and M. Trajtenberg (2002). *Patents, Citations & Innovations: A Window on the Knowledge Economy*. MIT press.

Jeong, S., J. C. Kim, and J. Y. Choi (2015). "Technology convergence: What developmental stage are we in?" In: *Scientometrics* 104.3, pp. 841–871.

Johnstone, N., I. Haščič, and D. Popp (2010). "Renewable energy policies and technological innovation: Evidence based on patent counts". In: *Environmental and Resource Economics*.

Joo, S. H. and Y. Kim (2010). "Measuring relatedness between technological fields". In: *Scientometrics* 83.2, pp. 435–454.

Kogler, D. F., D. L. Rigby, and I. Tucker (2013). "Mapping Knowledge Space and Technological Relatedness in US Cities". In: *European Planning Studies* 21.9, pp. 1374–1391.

Krafft, J., F. Quatraro, and P. P. Saviotti (2011). "The knowledge-base evolution in biotechnology: A social network analysis". In: *Economics of Innovation and New Technology* 20.5, pp. 445–475.

Lanoie, P., J. Laurent-Lucchetti, N. Johnstone, and S. Ambec (2011). "Environmental policy, innovation and performance: new insights on the Porter hypothesis". In: *Journal of Economics & Management Strategy* 20.3, pp. 803–842.

Laursen, K. and A. Salter (2006). "Open for innovation: The role of openness in explaining innovation performance among U.K. manufacturing firms". In: *Strategic Management Journal* 27.2, pp. 131–150.

Leten, B., R. Belderbos, and B. Van Looy (2007). "Technological diversification, coherence, and performance of firms". In: *Journal of Product Innovation Management* 24.6, pp. 567–579.

Leydesdorff, L. (2008). "Patent classifications as indicators of intellectual organization". In: *Journal of the American Society for Information Science and Technology* 59.10, pp. 1582–1597.

Leydesdorff, L. and I. Rafols (2009). "A global map of science based on the ISI subject categories". In: *Journal of the American Society for Information Science and Technology* 60.2, pp. 348–362.

Nelson, R. R. and S. G. Winter (1982). "An evolutionary theory of economic change". In: *Harvard Business School Press, Cambridge.*

Nesta, L. and P. P. Saviotti (2005). "Coherence of the knowledge base and the firm's innovative performance: Evidence from the U.S. pharmaceutical industry". In: *Journal of Industrial Economics* 53.1, pp. 123–142.

Nooteboom, B. (2000). *Learning and Innovation in Organizations and Economies.* Vol. 47. 4. Oxford University Press, p. 752.

Olsson, O. (2000). "Knowledge as a set in idea space: An epistemological view on growth". In: *Journal of Economic Growth* 5.3, pp. 253–275.

Oltra, V. and M. S. Jean (2005). "Environmental innovation and clean technology: an evolutionary framework". In: *International Journal of Sustainable Development* 8.3, p. 153.

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

Pavitt, K. (1985). "Patent statistics as indicators of innovative activities: Possibilities and problems". In: *Scientometrics* 7.1-2, pp. 77–99.

Petruzzelli, A. M., R. M. Dangelico, D. Rotolo, and V. Albino (2011). “Organizational factors and technological features in the development of green innovations: Evidence from patent analysis”. In: *Innovation: Management, Policy and Practice* 13.3.

Popp, D. (2003). “Pollution control innovations and the Clean Air Act of 1990”. In: *Journal of Policy Analysis and Management* 22.4, pp. 641–660.

Porter, M. E. and C. v. d. Linde (2011). “Toward a New Conception of the Environment-Competitiveness Relationship”. In: *Journal of Economic Perspectives* 9.4, pp. 97–118.

Quatraro, F. (2010). “Knowledge coherence, variety and economic growth: Manufacturing evidence from Italian regions”. In: *Research Policy* 39.10, pp. 1289–1302.

Quatraro, F. and A. Scandura (2019). “Academic Inventors and the Antecedents of Green Technologies. A Regional Analysis of Italian Patent Data”. In: *Ecological Economics* 156, pp. 247–263.

Rafols, I. (2014). “Knowledge integration and diffusion: Measures and mapping of diversity and coherence”. In: *Measuring scholarly impact*. Springer. Chap. 8, pp. 169–190.

Rafols, I. and M. Meyer (2010). “Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience”. In: *Scientometrics* 82.2, pp. 263–287.

Rao, C. R. (1982). “Diversity and dissimilarity coefficients: A unified approach”. In: *Theoretical Population Biology* 21.1, pp. 24–43.

Schmoch, U. (2008). *Conception of a Technology Classification for Country Comparisons*. Tech. rep.

Schumpeter, J. (1934). “The theory of economic development”. In: *Joseph Alois Schumpeter*.

Shannon, C. E. (1948). “A mathematical theory of communication”. In: *The Bell System Technical Journal*.

Stirling, A. (2007). “A general framework for analysing diversity in science, technology and society”. In: *Journal of the Royal Society Interface* 4.15, pp. 707–719.

Strumsky, D., J. Lobo, and S. van der Leeuw (2012). “Using patent technology codes to study technological change”. In: *Economics of Innovation and New Technology* 21.3, pp. 267–286.

Teece, D. J., R. Rumelt, G. Dosi, and S. Winter (1994). “Understanding corporate coherence. Theory and evidence”. In: *Journal of Economic Behavior and Organization* 23.1, pp. 1–30.

Uzzi, B., S. Mukherjee, M. Stringer, and B. Jones (2013). "Atypical combinations and scientific impact". In: *Science* 342.6157, pp. 468–472.
Wallace, M. L., Y. Gingras, and R. Duhon (2009). "A new approach for detecting scientific specialties from raw cocitation networks". In: *Journal of the American Society for Information Science and Technology* 60.2, pp. 240–246.
Yan, B. and J. Luo (2017). "Measuring technological distance for patent mapping". In: *Journal of the Association for Information Science and Technology* 68.2, pp. 423–437.
Zeppini, P. and J. C. van Den Bergh (2011). "Competing recombinant technologies for environmental innovation: Extending arthur's model of lock-in". In: *Industry and Innovation* 18.3, pp. 317–334.

Appendix A - Illustrative Patents Example

Espacenet

Bibliographic data: EP2551856 (A1) — 2013-01-30

HIGH FREQUENCY CABLE AND HIGH FREQUENCY COIL

Inventor(s):	SHINMOTO TAKASHI [JP]; YOSHIDA SHOTARO [JP]; KATAYAMA SHINJI [JP]; TODA TAIKOU [JP]; KATO TAKAMASA [JP]; DAIBO MASANORI [JP]; KAWAKAMI AKIO [JP] + (SHINMOTO, TAKASHI, ; YOSHIDA, SHOTARO, ; KATAYAMA, SHINJI, ; TODA, TAIKOU, ; KATO, TAKAMASA, ; DAIBO, MASANORI, ; KAWAKAMI, AKIO)
Applicant(s):	FUJIKURA LTD [JP] + (FUJIKURA LTD)
Classification:	**- international:** ***H01B5/02; H01B7/30; H01F27/28*** **- cooperative: B21C3/04 (US); B23K20/001 (US); B23K20/2333 (US); B32B15/01 (EP, US); C22C21/00 (EP, US); C22C9/00 (EP, US); C23C28/023 (EP, US); H01B1/023 (EP, US); H01B1/026 (EP, US); H01B13/0006 (US); H01B5/02 (EP, US); H01B7/30 (EP, US); H01F27/2823 (EP, US);** B21C1/003 (US); Y10T29/49117 (EP, US)
Application number:	EP20110759439 20110323 Global Dossier
Priority number(s):	JP20100066793 20100323 ; WO2011JP56984 20110323
Also published as:	EP2551856 (A4) EP2551856 (B1) CN102822907 (A) CN102822907 (B) JP4879373 (B2) more

Abstract of EP2551856 (A1)

A high frequency cable includes: a central conductor 1 made from aluminum or an aluminum alloy; a covering layer 2 made from copper covering the central conductor 1, and having a fiber-like structure in a longitudinal direction; and an intermetallic compound layer 3 formed between the central conductor 1 and the covering layer 2 and having greater volume resistivity than the covering layer 2, wherein a cross-sectional area of the covering layer 2 is 15% or less of an entire cross-sectional area including the central conductor 1, the intermetallic compound layer 3 and the covering layer 2.

Figure A1: Front page of the "High Frequency Cable" patent document. Source: `https://worldwide.espacenet.com/`

Espacenet

Bibliographic data: EP2598400 (A2) — 2013-06-05

HIGH ALTITUDE PLATFORM

Inventor(s): DAVIDSON PETER [GB]; HUNT HUGH EDMUND MURRAY [GB]; BURGOYNE CHRISTOPHER JOHN [GB] + (DAVIDSON, PETER, ; HUNT, HUGH EDMUND MURRAY, ; BURGOYNE, CHRISTOPHER JOHN)

Applicant(s): DAVIDSON TECHNOLOGY LTD [GB] + (DAVIDSON TECHNOLOGY LIMITED)

Classification:
- **international:** ***B64B1/50; B64C39/02; D07B1/14; D07B5/00; H01M8/22; H02K7/18***
- **cooperative:** **B64B1/44 (GB); B64B1/50 (EP, GB, US); B64C39/022 (EP, GB, US); D07B1/14 (EP, US); D07B1/147 (GB); D07B1/148 (GB); F02C6/00 (US); H01M8/04201 (US); H01M8/22 (EP, US); H02K7/1807 (EP, US); H04B10/25752 (US); H04B7/18502 (GB); H04B7/18504 (GB, US);** D07B5/005 (EP, US); H01M2250/20 (US); Y02T90/32 (EP)

Application number: EP20110735903 20110614 Global Dossier

Priority number(s): GB20100017685 20101020 ; GB20100015807 20100921 ; GB20100012864 20100730 ; WO2011GB51109 20110614

Also published as: EP2598400 (B1) AU2011284476 (A1) AU2011284476 (B2) CA2803958 (A1) CN103180206 (A) more

Abstract not available for EP2598400 (A2)
Abstract of corresponding document: GB2482340 (A)

A high altitude platform 4 is supported by a balloon 3 and is anchored by a tether 2 to a location at ground level such as a ship 1. The platform may include apparatus for generating electrical energy at altitude such as a fuel cell. The tether may include a conduit 6 for transferring a fluid fuel to the generating unit. The

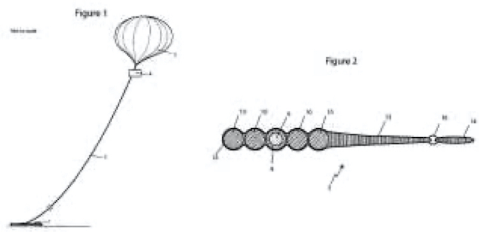

Figure A2: Front page of the "High Altitude Platform" patent document. Source: `https://worldwide.espacenet.com/`