Barbieri, N.; Perruchas, FDX.; Consoli, D. (2020). Specialization, Diversification, and Environmental Technology Life Cycle. Economic Geography. 96(2):161-186. https://doi.org/10.1080/00130095.2020.1721279

Specialization, diversification and environmental technology life-cycle

Forthcoming in *Economic Geography*

Nicoló Barbieri

Universitá di Ferrara (Italy)

nicolo.barbieri@unife.it

François Perruchas

INGENIO (CSIC-UPV), Universitat Politècnica de València, Valencia (Spain)

franperr@ingenio.upv.es

Davide Consoli

INGENIO (CSIC-UPV), Universitat Politècnica de València, Valencia (Spain)

davide.consoli@ingenio.upv.es

Abstract. The paper analyses whether and to what extent regional related and unrelated variety matter for the development of green technology, and whether their influence differs over the technology life-cycle. Using patent and socio-economic data on a thirty-year (1980-2009) panel of US States, we find that unrelated variety is a positive predictor of green innovative activities. When unpacked over the life-cycle, unrelated variety is the main driver of green technology development in early stages while related variety becomes more prominent as the technology enters into maturity.

KEYWORDS: Green technology; Technology life-cycle; Regional Diversification

JEL: O33; Q55; R11

1 Introduction

The debate on regions' ability to create and attract new industries, as well as to managing their decline, is a staple of economic geography. Inspired by Schumpeter's (1939) tenet of creative destruction, scholars concur that innovation is the main thrust of regional development as industries move from infancy to maturity and, eventually, decline (Norton and Rees, 1979; Gort and Klepper, 1982; Markusen, 1985; Storper and Walker, 1989; Klepper, 1997; Audretsch and Feldman, 1996). Other studies point out that, akin industries, also technologies develop along a life-cycle path (e.g., Abernathy and Utterback, 1978; Abernathy and Clark, 1985; Achilladelis et al., 1990; Achilladelis, 1993; Andersen, 1999). Although these works explore primarily the characteristics of technology evolution (e.g. take-off, duration of the stages, etc.), the question of 'where' different technologies develop along the life-cycle, and how this connects with the local competence base has been neglected. The present paper fills this gap by investigating whether and to what extent regional knowledge structures matter for the development of technology, and whether their influence differs along the path of technology life-cycle. We do so by focussing on green technology, a particular instantiation of innovation consisting of standards and artefacts aimed at mitigating or reversing the negative effects of human action on the environment.

The contributions of the present paper are manifold. First, we engage the literature on the spatial contingencies along the industry life-cycle (Henderson et al., 1995; Duranton and Puga, 2001; Neffke et al., 2011a). Therein, learning opportunities provided by local diversity attract new, young industries, whereas mature ones thrive in specialised local environments. While prior research focuses on industry evolution, we focus on the technology life-cycle with a view to shed light on the association between different structures of regional know-how and technological progress. To our knowledge, this is the first attempt to operationalise the empirical connection between the technology life-cycle and the knowledge base, which had so far only been approached on conceptual grounds (Vona and Consoli, 2015). In so doing, the paper draws on the long-term perspective on the evolution of the regional knowledge base to systematically capture the diversity of technology dynamics.

Second, we take the cue from literature on the effects of regional characteristics along the industry life cycle arguing that agglomeration economies are not dichotomous. Rather, accounting for variation of the regional knowledge base over degrees of relatedness affords the opportunity to better identify the drivers of knowledge generation. Economic geographers and innovation scholars maintain that the more diverse the spectrum of know-how in a region, the greater the potential of successfully exploiting both available inputs and unexplored interdependences between them (Rigby and Essletzbichler, 1997; Frenken and Boschma, 2007; Balland and Rigby, 2016). This rests on the premise that the composition of activities through which knowledge is channelled into productive uses affects the rate and direction of technical change in a region. In this vein, the more sectors are related, the easier is knowledge transfer from one domain of application to another. A review of empirical studies by Content and Frenken (2016) confirms that relatedness is an important driver of regional diversification across various dimensions (e.g., products, industries, technologies) and spatial units (e.g., countries, regions, cities, labour market areas), but also concludes that the evidence is still mixed. Related diversification is often found to be a stronger driver compared to unrelated diversification, not surprising considering the nature of these constructs. Diversification is an uncertain process that can be better dealt with by relying on available local resources, and on well-tested connections across them, both trademark features of related variety. Unrelated diversification, on the other hand, entails implementing new forms of coordination across different and formerly unassociated capabilities (Desrochers and Leppälä, 2011; Boschma, 2017). At the same time, Boschma and Frenken (2006) call for caution against determinism, highlighting that spatial contingencies are less important at early stages of a sector's development due to gaps between the requirements of new knowledge and the established environment.

Within this debate, the question of whether and to what extent both related and unrelated variety affect technological innovation was addressed only recently by Castaldi et al. (2015). Their empirical analysis on the United States (US) shows that the two forms of regional diversification are not opposite but, rather, complementary forces. In particular, radical innovation is more frequent in federal states with a diversified knowledge base across unrelated domains, whereas incremental innovation has a stronger association with related variety in local knowledge. The present paper contributes to this strand by distinguishing between related and unrelated variety along the life cycle path of green technology. In so doing, we take issue with the notion that either related or unrelated variety are drivers of innovation regardless of the life-cycle stage of the technology.

Third, we propose that it is important to consider simultaneously region-specific and external factors that may favour the emergence of new technologies. To this end, we adopt a regional knowledge production (RKP) function approach that incorporates qualitative features of the local knowledge base as well as the life cycle of technology. So far, the literature has looked at the extent to which R&D and human capital interact (Charlot et al., 2015) and affect (Crescenzi et al., 2015) regional innovation. However, following the evolutionary tenet that innovation stems from recombination of existing ideas (Schumpeter, 1939; Basalla, 1989; Weitzman, 1998; Arthur, 2007), our contribution extends the RKP framework to account for the influence of regional knowledge bases and of relatedness between their components (Frenken et al., 2007; Castaldi et al., 2015). Against this backdrop, we expect that the particular stage of a technology's life cycle determines whether local diversification across knowledge domains benefits innovation.

Fourth, we provide new empirical evidence on the connection between environmental sustainability and regional studies. From a policy perspective, the green economy is often touted as holding the potential for new growth and job creation. At the local scale, the pressure is on regions' and countries' institutions to create the adequate premises for innovation in adaptation and mitigation strategies. A pillar of the present paper is that green technologies, usually treated as a homogenous block, differ substantially from one another. The paper claims novelty in being the first to empirically grasp one dimension of this heterogeneity, namely the degree of technology maturity, which we measure as a combination of volume of inventive efforts and of geographical distribution. In so doing, it also pushes the agenda of the nascent, but still underdeveloped (Truffer and Coenen, 2012) sub-discipline of environmental economic geography.

The study builds on the above conceptual grounds to test two conjectures. The first is that unrelated variety of the local knowledge stock matters for innovation at early stages of the technology life cycle, while related variety has little or no effect. The second is that, as technology moves towards maturity, related variety is the main driver and unrelated variety progressively loses prominence. The empirical analysis is on green technology development in a panel of 48 US federal states and District of Columbia (D.C.) between 1980 and 2009. Our main data source is the catalogue of patent applications in PATSTAT from which we extract information on patent families to develop an original indicator for the stage of development of green technologies and on the location of inventors. For the goal of studying the relationship between technology life cycle and regional knowledge structure, we build entropy indicators at different levels of relatedness between technological domains (Jacquemin and Berry, 1979; Attaran, 1986; Frenken et al., 2007;

Castaldi et al., 2015). Finally, we follow the parametric approach of Charlot et al. (2015) and adopt a random growth specification of the unobservable part of the model to control for time-invariant regional characteristics, common time effects and time-varying unobservable features.

The analysis yields two main findings. First, green technology development exhibits stronger association with unrelated variety than with related variety. This is for two reasons. On the one hand, the transition towards environmentally sustainable production is, on the whole, at early stages (OECD, 2015). On the other hand, green technology is more complex than non-green technology and therefore requires the orchestration of diverse, and cognitively distant, knowledge inputs (De Marchi, 2012; Cainelli et al., 2015; Barbieri et al., 2018). The second key finding, in line with our expectation, is that unrelated variety has stronger association with early stages of the green technology life cycle while related variety becomes more important as technology enters into maturity.

The remainder of the paper is organised as follow. The next section overviews the relevant literature. Section 3 outlies the data, variables and empirical strategy. Section 4 presents the descriptive statistics and discuss the results. Section 5 concludes and summarises.

2 Theoretical background

2.1 Industry life cycle and agglomeration economies

Research in economic geography points to two key mechanisms of regional development. The first builds on Marshall's (1920) tenet based on the interaction and proximity of goals and competences, whereas the second stems from Jacobs' (1969) emphasis on the diversity of competences in the local economy. Building on these insights, scholars have often pointed out that the benefits of agglomeration externalities depend on life cycle dynamics. Empirical evidence both from regional economics (Norton and Rees, 1979; Markusen, 1985) and industrial dynamics (e.g. Gort and Klepper, 1982; Abernathy and Clark, 1985; Storper and Walker, 1989; Audretsch and Feldman, 1996; Klepper, 1996; 1997; Agarwal and Gort, 2002) supports the conjecture that emerging industries grow at a faster pace than those locked into old, mature industries.[1] Duranton and Puga (2001)

[1] For instance, Norton and Rees (1979) find that the decline of the US Manufacturing Belt during the late sixties was essentially a core-periphery realignment, which has theoretical roots in the product life cycle framework. The decentralisation of production towards peripheral Southern and Western states followed the dispersion of innovative capacity and the rise of new, high-tech sectors at the beginning of the life cycle.

elaborate a conceptual framework to explains how diversification and specialisation favour, respectively, young and mature industries. At the beginning of the life cycle, when young firms need to experiment with prototypes of new products, diversified local environments are seedbeds for alternative production processes that can be tried, adopted or discarded. When the product design reaches maturity and are ripe for mass production specialised cities provide a more suitable environment due to benefits such as knowledge pools and lower production costs. Empirical studies corroborate this framework showing that there is a continuity between Marshall and Jacob once the life cycle is accounted for (Henderson et al., 1995; Neffke et al, 2011a) and that diversification facilitates regional growth due to knowledge spillovers and learning opportunities in diverse urban settings (Glaeser et al., 1992; Duranton and Puga, 2001; Frenken et al., 2007; Neffke et al., 2011b). Therein Jacob's externalities favour the adoption of new processes and products whereas at early stages Marshall externalities could even be detrimental (Harrison et al., 1996; Kelley and Helper, 1999; Feldman and Audretsch, 1999; Castaldi et al., 2015).

The conceptual explanation of the positive relationship between regional diversification and innovation is rooted in the recombinant innovation theory (Schumpeter, 1939; Nelson and Winter, 1982; Weitzman, 1998; Fleming, 2001) whereby higher variety of local know-how increases the likelihood of original recombination and innovation. In this context, local search and bounded rationality are important dimensions (March and Simon, 1958; Nelson and Winter 1982) in that innovators recombine bits of knowledge they are familiar with in order to decrease the risk of failure. In so doing, however, they reduce the chances of developing radical innovation. On the contrary, innovators who recombine cognitively distant bits of knowledge face higher uncertainty but, also, higher payoffs if the innovative effort is successful.

The recent evolutionary turn in economic geography builds on tenet that Jacobs externalities do not merely lead to a more efficient division of labour within regions. Rather, in a diversified environment the opportunities for innovation increase due to the availability of different types of knowledge that is geographically close and can be recombined. Along these lines, Frenken et al. (2007) shifted the debate on agglomeration economies by acknowledging that diversification per se does not fully capture the mechanism that spur regional economic growth. The flow of knowledge within regions requires a balance of cognitive distance to avoid lock-ins and of cognitive proximity to enable effective learning (Nooteboom, 2002; Boschma and Iammarino, 2009). Such an intuition has given way to the notion of related (unrelated) variety. In short, two industries are related when they share some cognitive structures that enhance learning opportunities and knowledge spillovers. Relatedness, in turn, has been to yield benefits such as regional growth (Frenken et al., 2007; Essletzbichler 2007; Bishop and Gripaios 2010). A recent review by Content and Frenken (2016) however concludes that evidence on the prominence of related variety is still mixed, and more research should articulate the role of unrelated variety. This echoes Boschma and Frenken's (2006) warning over the risk that sectoral and territorial dynamics change, and so do the connections with the underlying knowledge base.

Adding to the latter, we observe that prior literature portrays related and unrelated variety as binary, mutually exclusive, drivers. In a first attempt to debunk this view Castaldi et al. (2015) show that radical innovations emerge in regions with diversified knowledge bases – i.e. connecting distant technological domains – whereas incremental innovation is more likely in regions that feature related variety in local knowledge. In the following subsection we will build on this insight.

2.2 Technology life cycle in the regional knowledge production function

The literature reviewed so far emphasises the pivotal role of technological change in regional development. Along the life cycle industries rely on different types of innovation that require different sources (Norton and Rees, 1979). The birth of new industries typically follows radical innovation and the development of immature technologies, whereas once a dominant design is established, technological disruptions are less likely and the industry reaches a maturity stage in which innovation is mostly incremental (Neffke et al., 2011a). Such a mechanism implies that industries exploit different types of agglomeration externalities according to their stage of maturity. Existing studies however treat technology as a latent element that evolves and leads to industry maturity. Thereby, agglomeration economies are beneficial for industry and regional growth because of their indirect effect via knowledge spillovers and learning opportunities. No study has, to this day, directly studied the mechanisms through which agglomeration externalities trigger technology.

Like industries or products, technology evolves along a S-shaped (or double S-shaped) life cycle path: introduction, growth, maturity and decline (Achilladelis et al., 1990; Achilladelis, 1993; Andersen, 1999; Haupt et al., 2007). In the earliest phase different pieces of knowledge are recombined to obtain a new technology that differs from what was before. Herein, usually few firms experiment and strive with the high degree of uncertainty. The technology that emerges from this phase is often associated with high production costs, low penetration in the market and uncertainty in the potential use of the technology itself (Callon, 1998). In the subsequent phase, uncertainty is lower, risk associated to R&D decrease, innovation is less radical and the number of competitors increases (Haupt et al., 2007). Finally, when a dominant design is reached the technology enters a maturity phase that mainly advances through incremental innovation with high degrees of standardisation and widespread diffusion.

A critical issue in the diffusion literature is the implicit assumption that neither the new technology nor the one that is being replaced change (Hall, 2004). This static view stands in contrast with empirical evidence on the incremental adaptations that ultimately leads to improvement of technology (Christensen, 1997; Foster, 1986). Moreover, and closer to our goal, the balance between intrinsic performance characteristics and the specific features of the selection environment is central to the dynamics technology (Vona and Consoli, 2015). These features can be bottlenecks – see e.g. the analysis of the American machine tool industry by Rosenberg (1976) or Hughes' (1983) account of the evolution of the electrical power system – or can be facilitating circumstances of the ecosystem – as is the case in Constant's (1980) study on aircraft piston-engine or in Henderson's (1995) analysis on optical lithography. The broader point is that acknowledging the role of the context of adoption entails shifting the focus from substitution between new and old technology to the evolution of the selection environment. This resonates with Boschma and Frenken's (2006) cautionary remark about deterministic accounts of regional variety: spatial contingencies, and the associated uncertainties, matter.

Building on these premises, we study how agglomeration economies and technology life cycle interact. In the geography of innovation literature, the RKP function approach provides a suitable theoretical framework to investigate these issues (see e.g. Crescenzi et al., 2007, 2012; Ponds et al., 2010; Feldman et al., 2014; Charlot et al., 2015). Therein the regional perspective is embedded in the knowledge production function framework proposed by Griliches (1979) to observe the regional determinants of the generation of innovation. However, whether regional innovation inputs (e.g. human capital and R&D investments) and agglomeration economies exert heterogeneous effects on innovation output according to the maturity of the technology remains an unexplored question. We expect that delving into these details will provide useful insights into the type of know-how that enables regions to innovate and move along the life cycle. Accordingly, we extend the RKP framework to incorporate knowledge diversification at different levels of variety (Frenken et al., 2007; Castaldi et al., 2015). Moreover, since the regional endowment of innovation inputs and given the heterogeneity of regional structural characteristics, we also test whether specific features of the local knowledge base exert different impacts on innovation output depending on the level of development of regions.

2.3 Green technological developments

The present paper focuses on a particular instantiation of innovation that aims at reducing the impact of human activity on natural resources and ecosystems. Environmental-related innovation attracts growing interest in the economics of innovation (see reviews by Popp et al., 2010; Barbieri et al. 2016). The literature points out that the transition to low-carbon economies does not happen on a blank canvas, and that achieving green growth entails dealing with the inertia of existing productive structures which ought to be adapted or dismantled while new ones are put in place. Such a path entails striking a delicate balance between achieving lower environmental impacts while maintaining efficiency (Ghisetti et al., 2015). This higher complexity calls upon a broader and diverse set of skills, knowledge inputs and competences (De Marchi, 2012; Vona et al, 2018).

Recent evidence also indicates that green technologies exhibit peculiar geographical connotations (Truffer and Coenen, 2012). The literature emphasises that the development of green technologies is crucial to sustainable local development (Montresor and Quatraro, 2019). Spatial local levers, such regional knowledge spillovers and agglomeration economies, facilitate the development and the adoption of eco-innovations (Cainelli et al. 2012; Antonioli et al., 2016). Even in the case of green technologies, diversification and relatedness are key to green technology and industry development. Using patent data, Tanner (2014) and Montresor and Quatraro (2019) show that high endowment of local environmental related knowledge is a positive predictor of green technological development. We contribute to this literature by articulating the connection between local eco-innovation capacity and degrees of knowledge relatedness.

3 Empirical application

3.1 Data

The empirical analysis builds on an original dataset that incorporates information on patenting activities and socio-economic data in 49 US Federal States over the period 1980-2009. The use of patent data in economic geography has gained relevance recently (see e.g. Crescenzi et al., 2007; Tanner, 2014; Balland and Rigby, 2017; Montresor and Quatraro, 2017). Patents provide a wealth of information on inventive activities such as the location of the inventors or applicants, the knowledge base of the invention (such as citations to prior patents or studies), and its technical content. A key benefit of this type of data is the granular analysis of specific knowledge domains (Popp, 2005), which is essential for the purposes of the present paper. At the same time, use of patent data entails some limitations. For instance, not all the inventions are patented and the quality of patents varies according to the technology under study (Hascic & Mingotto, 2015). Nevertheless, since patents are usually filed early in the innovation process, they provide a good indicator of research and development activities (Griliches, 1990).

Patent data are extracted from the 2016 version of PATSTAT (source: European Patent Office, EPO). Relevant to our analysis is the subset of environmental-related patents identified through the Env-Tech classification of the OECD (2016), which lists International Patent Classification (IPC) and Cooperative Patent Classification (CPC)[2] codes concerning 95 environmental-related technologies, grouped into 8 families and 36 subgroups.[3] These technologies are designed to reduce anthropogenic pressure on natural resources and improve adaptation to the changing environment and, as such, encompass a broad spectrum of domains including environmental pollution, water scarcity and climate change mitigation (Hascic & Mingotto, 2015).[4] We collected all environmental-related patents coded in the Env-Tech classification and, following prior literature, extracted from PATSTAT information on patent families, which are our main unit of

[2] Patent offices use IPC and CPC to classify patent documents. Both classification systems exhibit a hierarchical structure based on the technical content of the patents through codes. At the lowest level, i.e. full-digit, the codes are very specific and refer to narrow technological fields, e.g. IPC full-digit G06F9/02 – "Arrangements for program control using wired connections". At the highest level, i.e. 1-digit, the codes refer to broad technological domains, e.g. IPC 1-digit G - "Physics".

[3] In an intermediate step, we convert the IPC codes listed in the Env-Tech into CPC codes using a correspondence table provided by the European Patent Office (EPO) and the United States Patent and Trademark Office (USPTO). This allows us to use a unique classification system.

[4] Specifically, Env-Tech classification codes cover eight technology groups: environmental management, water management, energy production, capture and storage of greenhouse gases, transportation, buildings, waste management and production of goods.

analysis (see e.g. Hall and Helmers, 2013). To avoid double counting of inventions for which protection was sought at different national offices, we identify 1,071,869 environmental-related patent families filled between 1980 and 2009.

3.2 Measuring regional knowledge base

To measure the regional knowledge base, we use information from PATSTAT and assign patent families to US states by feeding the geographical coordinates of inventors' addresses into GeoNames[5], a database containing worldwide geographical information on, among others, administrative borders and postal codes.[6] In order to reduce noise, we limit the search to cities with at least five thousand inhabitants and manually check if the city name is far from the end of the address string. Finally, we use the Google Maps API – a programmable interface developed by Google since 2005, to assign the geographical coordinates of the remaining addresses not found in the first steps.[7]

In spite of EPO's constant updates, a non-negligible share of inventor's addresses is still missing from the PATSTAT database. Therefore, after the data cleaning process (detailed in Appendix A) we exploit the work by the Institut Francilien Recherche Innovation Société (IFRIS) where missing addresses have been filled using sources such as REGPAT and National Patent Databases.[8] This allows us to geo-localise 798,455 (74.5%) green patent families worldwide, 149,161 of which in the US (91.3 % have half or more of their inventors geo-localised and 67.1 % have all their inventors geo-localised). We then group patent families according to the state of residence of the inventor.

The choice of US federal states as units of analysis is dictated by data availability. Information for some key variables that are described below (e.g. R&D expenditures and human capital endowment) limits our analysis to this geographical level. Indeed, the long-

---

[5] The GeoNames database provides geographical coordinates of a wide range of features such as mountains, lakes, countries borders, etc. In particular, it includes information on the latitude and longitude of the majority of the cities around the world. See http://www.geonames.org (last access: 25 August 2019) for more information.

[6] When the postcode is missing, we identify the city in the address string using GeoNames, that is, we split addresses in several elements in order to isolate the street, city, etc. Then, since the city is usually provided at the end of the address, we browse the address string from right to left. Our algorithm compares each element of the address with the city name information included in GeoNames. We repeat this process for all the elements of the address string moving from the end to the beginning and associate the address to the city name in case of matching. For example, in the address: *John Smith, 1 West 72nd Street, New York, NY,* there are four elements to check: "John Smith", "1 West 72nd Street", "New York" and "NY". Starting from the right, the city will be detected in the second loop of the algorithm, i.e. New York.

[7] Daily search limits and costs did not enable us to use Google Maps API to search for the geographical coordinates of all addresses.

[8] For more details, check https://github.com/cortext/patstat (last access: 25 August 2019).

term perspective of our analysis (1980-2009) further restricts the availability of these information which are fundamental to control for in a regional knowledge production function. In particular, the state level enables us to control for idiosyncratic features of each state, such as environmental policy.

Figure 1 shows the geographical distribution of patenting activities at state level. Not surprisingly, most innovative states such as New York, Pennsylvania, Ohio, Illinois, California, Florida, Georgia and Texas, fall in the top quintile of the patent distribution. Moreover, Figure 1 also provides information on the percentage of green patents in each US state. It is worth noting that the greening process follows two main patterns. First, among the states in the top quintile of total patenting the share of green patents ranges from less than 5% in California to a 7.6% in Pennsylvania. Second, another noticeable element is that some states rank high in terms of percentage of green patent families and low in total patenting activities, i.e. Wyoming (10.7%), West Virginia (8.1%) and Main (8.5%). This is probably due to the patenting history of these states.

The overall green patenting trend by groups as per Env-Tech (OECD, 2016) is reported in Figure 2. Therein we observe that patenting in most green technologies experience an acceleration after 2000. Technologies that improve the sustainability of the energy and building sectors lead the trend followed by green products and processes and transportation. The number of patent families related to water-related and carbon capture and storage technologies is relatively smaller compared to other green technologies (Panel A). However, the latter increases faster relative to 1980 levels, as showed in Panel B. Patenting on transportation and energy efficiency buildings experiences a sharp increase after 2005. Conversely, environmental management and water-related technologies exhibit lower growth rates over the period.

FIGURES ONE AND TWO ABOUT HERE

3.3 Measuring regional knowledge base diversification

To measure diversification of regional innovative activities we calculate entropy indicators that can be scaled at different levels of aggregation associated with specific degrees of relatedness. In the seminal paper by Frenken et al. (2007), the entropy measure is decomposed into related and unrelated variety to capture the extent to which relatedness and diversification characterise the regional cognitive structures. Castaldi et al. (2015)

employ the same measure to assess diversification in technological capabilities of US federal states. In the present paper, we follow Castaldi et al. (2015) in the use of geographical information on patent families (as detailed in Section 3.2) to calculate the entropy indicators using patent data at the state level in US. To do so, we exploit the technological classification codes assigned to each patent. The hierarchical structure of the International Patent Classification (IPC) system can be exploited to calculate variety at different code digits.[9] In the present paper the IPC codes are used to capture the technological fields that characterise the inventions.

We calculate related, semi-related and unrelated variety of invention activities assuming relatedness between two patents when they share an IPC code at different level of disaggregation. Accordingly, relatedness increases together with the number of IPC digits. Specifically, unrelated variety (UV) is measured using the entropy of the patent family distribution over IPC 3-digit classes:

$$UV_{it} = \sum_{k=1}^{N} pf_{k,it} \ln\left(\frac{1}{pf_{k,it}}\right)$$

Where $pf_{k,it}$ is the share of patent families in technological section $k = [1 \ldots N]$ at IPC 3-digit level, with at least one inventor located in state i at time t. Semi-related variety (SRV) is equal to the entropy at 4-digit within each IPC 3-digit section. Given the decomposition theorem developed by Theil (1972), SRV is the difference between the entropy measure calculated at 4-digit and 3-digit level (i.e. UV):

$$SRV_{it} = \sum_{l=1}^{P} pf_{l,it} \ln\left(\frac{1}{pf_{l,it}}\right) - UV_{it}$$

Where $pf_{l,it}$ represents the share of patent families in each state over technological subclasses $l=[1 \ldots P]$ (IPC 4-digit level). Finally, we calculate related variety (RV) at the IPC 8-digit level (subgroups). As before, RV is obtained by subtracting to the entropy at 8-digit, the one at 4-digit level (subclass). In so doing, we calculate variety across narrow

[9] IPC 3-digit classes capture generic domains of application while higher disaggregation, IPC 8 digits, refer to specific applications. The first 4 digits of the code indicate the class and subclass, whereas the last 4 digits are the groups and subgroups. To illustrate, IPC code "A61B 5/022" identifies inventions that allow the "measurement of pressure in heart or blood vessels by applying pressure to close blood vessels", while subclass "A61B" describes inventions related to "diagnosis, surgery and identification", and finally class "A61" is associated with "Medical or Veterinary Science; Hygiene".

technological subgroups (i.e. IPC 8-digit level) within each broader technological subclass (i.e. 4-digit level):

$$RV_{it} = \sum_{m=1}^{R} pf_{m,it} \ln\left(\frac{1}{pf_{m,it}}\right) - \sum_{l=1}^{P} pf_{l,it} \ln\left(\frac{1}{pf_{l,it}}\right)$$

Where $pf_{m,it}$ is the share of patent families in state i at time t over technological subgroups $m=[1...R]$. As far as we move from UV to RV, the cognitive distance between technological fields decreases. RV is calculated across very similar and specific technological fields (subgroups) compared to UV, which is measured across distant and broad technological fields (classes).

3.4 Measuring life cycle stages

To identify the maturity of green technologies, we develop a measure of technology life cycle based on two indicators: the geographical ubiquity of patenting and volume of patenting intensity. We calculate these using worldwide patent families for each macro-technology reported in the Env-Tech classification.[10] It is worth noting that such an exercise uses information on all patent families worldwide, and not only those filed in the US. This enables us to measure the overall stage of development of green technologies to which all worldwide inventors contributed to.

The ubiquity indicator captures the extent to which innovative activities are geographically spread relative to countries' specialisation in green technologies. Following Balland and Rigby (2017), the geographical scope of inventions is calculated using the Revealed Technological Advantage (RTA) for each green technology, country and time period as follows:

$$RTA_{cjt} = \frac{Patents_{cjt} / \sum_j Patents_{cjt}}{\sum_c Patents_{cjt} \,/\, \sum_{cj} Patents_{cjt}}$$

The RTA measures the intensity of the contribution of each country c to the development of Env-Tech technology j at time t. That is, it captures a country's efforts towards

[10] The Env-Tech classification OECD (2016) groups green technologies at different digits (up to four). In the present paper we exploit the 2-digit level which is a compromise between narrow (three digits) and broad (1-digit) technological fields. Moreover, the 2-digit level guarantees coverage of all the technologies listed in the classification since some of them (e.g. 4 and 3-digit codes) are not provided for all the technologies. Finally, the 2-digit level ensure that all the classes of the Env-Tech classification have at least one patent family over the period 1980-2009. Table 2 provides the list of green technological domains employed to define technology life cycle stages.

developing a specific green technology (numerator) relative to global efforts (denominator). This allows us to identify green technological domains that show a relative high rate of patenting in some countries compared to the world average. The ubiquity indicator of each Env-Tech technological field is given by the number of countries with RTA higher than one in a particular green technology at time *t*:

$$UBIQUITY_{jt} = \sum_{c} M_{cjt}$$

Where $M_{cjt} = 1$ if $RTA_{cjt} > 1$. Since *UBIQUITY* increases with the number of countries specialised in the development of a particular green technology, we interpret this indicator as a proxy for the diffusion of green inventive activities. The advantage of such a measure with respect to other patent indicators of diffusion like i.e. citations, family size, et cetera is that it captures specialisation patterns in specific green technologies relative to their global counterparts.

We calculate a second indicator based on the number of patent families at country level as a proxy of global invention intensity in green technologies. Finally, we measure the average growth rate over four years of both patenting intensity and ubiquity to smooth the trends and capture their dynamics over time.

Using ubiquity and patenting intensity together, we define the life cycle stages of each Env-Tech technological domain at world level. Table 1 shows that the *emergence* phase exhibits a low level of technological diffusion and intensity. This is the lowest level of maturity of the technology whereby inventive activities are concentrated in few countries and the number of patents is relatively low. To determine the step into maturity we follow two (non-exclusive) strategies. The first is moving from emergence to the *development* phase in which technological advances are still geographically concentrated and patenting activity increases at a faster pace. Alternatively, technologies can be in the *diffusion* phase, characterised by a growing number of countries specialised in the green technology and low patenting intensity. Finally, in the *maturity* phase standardisation in the design and knowledge-related activities is achieved, both patenting intensity and geographical diffusion of inventive activities are relatively high. We argue that such an approach affords a dynamic view of technological evolution, in that not all stages are always achieved, and, coherent with the framework of Section 2.2, maturity may be an intermediate stage before the appearance of further developments.

Building on the above, we assign green technologies to a particular stage of development as follows: we standardise the indicators and define the low (high) values shown in Table 1 if the technology exhibits ubiquity or patenting intensity below (above) the average value.[11] It is worth noting that such a procedure enables the technology life cycle indicator to be contingent on both idiosyncratic features of the technology under analysis and on the stage of development of the other green technologies. Table 2 reports the results of this process showing the life cycle stages of green technology in 1980, 1990, 2000 and 2010. The results of this exercise resonate with insights from specialised literature and policy reports. To illustrate, "Air pollution abatement" (ENV-TECH 1.1), "Renewable energy generation" (ENV-TECH 4.1), etc., is found in the maturity stage since the 1980s. Conversely, "Environmental monitoring" (ENV-TECH 1.5) or "Rail transport" (ENV-TECH 6.2) remain in the emergence phase with respect to other green technologies. Table 2 also shows some technologies that move from emergence to maturity stages – i.e. "Energy efficiency in buildings" (ENV-TECH 7.2), "Wastewater treatment" (ENV-TECH 8.1). Importantly, to reach maturity a technology does not necessarily go through all the life cycle stages in that development (high patenting and low ubiquity) and diffusion (low patenting and high ubiquity) are alternative pathways to achieve maturity.[12]

TABLES ONE AND TWO ABOUT HERE

Finally, we obtain the regional green technological efforts at each stage of the technology life cycle as follow:

$$GP_{it}^{L} = \sum_{j} P_{ij(L)t}$$

$$for\ each\ L = [Emergence, Development, Diffusion, Maturity]$$

where the green patent families in state i and time t are summed according to the life cycle stage L of green technology j they belong to (see Table 2 and Figure A1). The resulting four variables capture the geographical distribution of green patenting activities in each stage of the technology life cycle.

[11] We employ standard scores to normalise the values: $\frac{X-\mu}{\sigma}$ where X are the values of patenting intensity or ubiquity, μ is the mean and σ the standard deviation.

[12] An exhaustive description of the yearly patterns is provided in Appendix B.

Figure 3 shows the distribution of population-weighted green patenting across US states per life cycle stages, i.e. GP_{it}^{L}. A comparison across the different panels of the figure shows persistence of leading states in the top quintile of all stages of the life cycle. These states are also characterised by medium-high green patenting activity. Other states are more effective in the production of green technological knowledge only in some stages of the life cycle. For example, Washington ranks high in the development of green technologies in the emerging and diffusion stage, whereas Alabama in the development and diffusion stages. Michigan is effective especially in the production of knowledge related to emerging, developing and mature green technologies but not in those in the diffusion phase.

FIGURE THREE ABOUT HERE

3.5 The empirical model

To test whether and what type of knowledge base diversification is associated with the generation of new environmental technologies, the paper employs a Knowledge Production Function (KPF) inspired approach previously formalised by Griliches (1979) that is extended in three directions. First, following Jaffe (1989) and Crescenzi et al. (2007) we exploit the geographical dimension of the dataset (in our case US states), rather than focussing on firms (Jaffe, 1986), as unit of analysis to investigate the spatial organisation of innovative activities. Second, we acknowledge that local knowledge diversification plays a pivotal role in the knowledge production process (Jacobs, 1969; Glaeser et al., 1992) and that various forms of variety are associated with different degrees of relatedness between technological domains (Frenken et al., 2007; Castaldi et al., 2015). Third, we integrate the technology life cycle heuristic into the KPF framework in order to assess which type of variety in the knowledge base is associated with knowledge production process at different the levels of technological maturity.

We estimate the following empirical model:

$$GP_{jt}^{L} = \beta_1 \mathrm{Variety}_{\mathrm{jt}} + \beta_2\, R\&D_{jt} + \beta_3\, HC_{jt} + \mathrm{Controls}_{\mathrm{jt}} + \tau_{\mathrm{j}} + \gamma_{\mathrm{t}} + \delta_{\mathrm{jt}} + e_{jt} \qquad (1)$$

where the dependent variable is the number of patent families in all green technologies and separately for green technologies at different stages of the technology life cycle (L)

in state *j* and year t[13]. *Variety* is a proxy for regional knowledge base diversification discussed above that includes *UV*, *SRV* and *RV*. *R&D* are research and development expenditures and *HC* human capital. In some specifications we also include a battery of controls that capture R&D and human capital in neighbouring states and population density (*Controls*).[14] We also include time fixed effects (γ_t), state fixed effects (τ_j) and region specific time trends that control for unobservable heterogeneity that varies linearly over time in each state (Charlot et al., 2015). The latter enables us to capture, among others, state-specific time patterns that we are not able to control for due to data availability, such as policy intervention, green fiscal reforms, etc. which are usually introduced at federal state level. Overall, these issues supports the choice of adopting a state-level perspective.[15] Finally, e_{jt} captures the residual variation.

Given the pivotal inducement of regulation in environmental innovation (see Popp et al., 2010 and Barbieri et al., 2016 for a survey) the analysis should control for environmental policy implementation. A suitable strategy however ought to account for a number of issues. First, as previously stressed, the state-level perspective enables us to control for the idiosyncratic state-specific features that vary over time (e.g. environmental policy efforts). Second, environmental regulatory efforts may depend on the pressure that governments exert through monitoring and enforcement activities (Brunnermeier and Cohen, 2003). Penalty threats lead to higher abatement expenditures to comply with regulation. To operationalise regulatory monitoring and enforcement, we rely on prior literature and use the number of environmental inspections over the total of establishments per US state and year (see e.g. Laplante and Rilstone, 1996). Data on the number of establishments and inspections are collected from, respectively, US Census Bureau (County Business Patterns dataset) and Environment Protection Agency (ECHO dataset). That said, we acknowledge that the number of inspections better captures

[13] Alternative patent indicators have been used to test the robustness of our results. The findings described in the following section hold even if we consider just patent families with at least one granted patent. Results are available upon request.
[14] Neighbouring states are those that share a border.
[15] In a set of ancillary regressions, we control for concentration of inventive activities in some areas within states. Although data availability restricts our analysis at the state level in order to control for R&D, human capital, neighbouring states and environmental policy, we estimate a model in which the units of analysis are the Core-Based Statistical Areas (CBSAs). In this model R&D, human capital and environmental policy proxies are at state level whereas inventive activity is at the CBSA which is a combination of metropolitan and micropolitan areas. The results – including fixed effects at the CBSA level estimations – confirm the main findings. In an additional specification we also estimate the model in Equation 1 restricted to inventions developed in metropolitan statistical areas within each state. Once again, the robustness of our finding is confirmed. The results of these alternative empirical strategies are available upon request.

perceived stringency and use energy intensity as additional proxy for environmental regulation. Following previous studies, environmental policy stringency can be implicitly measured through its effects on energy consumption (Fredriksson and Vollebergh, 2009). Given the regulatory efforts of both energy and environmental policies to reduce energy consumption in the US (Joskow, 2002), energy efficiency carries the benefit of capturing combined regulatory strategies that concern a broader spectrum including physical energy constraints, performance standards, energy taxes, etc. (Fredriksson and Vollebergh, 2009; Chang et al., 2018). We measure energy intensity as the total energy consumed (in Btu) per unit of GDP (Fredriksson and Vollebergh, 2009). Data on energy consumption are from the U.S. Energy Information Administration, whereas data on GDP from the US Bureau of Economic Analysis.[16]

Table 3 provides descriptive statistics of the variables used in the econometric analysis.

TABLE THREE ABOUT HERE

4 Econometric results

Before exploring the results of the econometric analysis, Figure 4 provides a graphical indication of a positive relationship between patenting activities and variety at different level of relatedness. As regards related variety, green and total patents follow an almost-overlapping pattern with the majority of inventions stemming from states with higher related variety. The distribution of patenting over quintiles of unrelated variety shows that this type of diversification is particularly relevant for the generation of green knowledge relative to all patents. At lower levels of unrelated variety, total patenting prevails over green patenting. Conversely, the growth of unrelated diversification in the regional knowledge is associated with higher patenting.

FIGURE FOUR ABOUT HERE

The econometric estimation of the model detailed in Section 3.5 (Table 4) confirms these initial indications. We propose two main specifications to check for differences between

[16] Energy intensity is measured using data on energy consumption by all the sectors of the US economy (i.e. residential, commercial, transportation and industrial sectors). The results are robust if we employ data on energy consumption by the industrial sector. Moreover, we adopt an alternative strategy that accounts for emission intensity (CO_2 emissions from fossil fuel consumption per unit of GDP). This proxy allows controlling for the use of fossil fuels not merely employed for energy generation (Mazzanti et al, 2015). Using this alternative strategy does not change the main results. Tables are available upon request.

green patent families and non-green patent families. Common to all models is that whereas UV and RV variety are positive and statistically significant in the case of green patents, SRV and RV are positively associated with non-green patenting. This suggests that green inventive activities emerge in states where the knowledge base is diversified across unrelated technological domains. Focusing on the specification in Column 5, an increase of 1% in UV is associated with an increase of 1.5% in green patenting activities. For instance, moving from a US state at the 25th percentile of the distribution of UV (Maryland) in 2005 to a US state at the 75th percentile (Florida), increases the number of green patents by 11.4%. On the other hand, non-green patenting activities proliferate in states characterised by semi-related and related diversification. In addition, when testing the difference between the coefficients in each respective specification, UV and RV are significantly different at 5% in Column 5 while for non-green patenting the null hypothesis of equality between SRV and RV coefficients is rejected.[17] This confirms that green technologies require diversification across both unrelated and related knowledge domains, and differ from non-green patenting that instead relies more on related diversification. This is in line with studies that emphasise the higher complexity of green innovation relative to non-green due to the recombination of more distant know-how (De Marchi, 2012; Cainelli et al., 2015; Barbieri et al, 2018). The result also resonates with empirical studies showing that hybrid green and non-green competences matter for the transition to low carbon (Barbieri and Consoli, 2019; Quatraro and Scandura, 2019). Finally, we observe that human capital is positive and slightly significant across almost all specifications while R&D expenditures is not in both the green and non-green RKP functions.

TABLE FOUR ABOUT HERE

Moving to the core of our analysis, Table 5 presents estimates of the model using green patent families at each stage of the technology life cycle as dependent variable. Therein, the coefficient of UV is statistically significant for emerging technologies, thus implying that diversification across unrelated technological fields favours green technologies in the emerging phase. An increase of 1% in UV is associated with 1.1% additional green patenting, that is, moving from the 25th to the 75th percentile of the UV distribution leads to an increase of 8.4% in green patenting. According to the recombinant innovation theory, in the early stage of the life cycle technological development benefits from the

[17] The null hypothesis is rejected at 5%

richness of cognitively distant bits of knowledge. Together with unrelated variety, R&D expenditures play a key role in this stage of technology evolution. In the subsequent stage of the life cycle (i.e. development), characterised by higher patenting intensity, all types of variety exert a positive effect on green innovative activities. In this phase, human capital is positively associated with green patent production. Moving to the diffusion phase, related variety in the local knowledge base is positively correlated with the generation of environmental-related patents, although the coefficient is significant at only 10%.[18] In addition, both the main innovation inputs, i.e. R&D and human capital are positive and significant. Finally, at maturity, related variety becomes the main driver of green innovative activities. In this case, a 1% increase in RV is associated with a 0.6% increase in green patenting so that moving from the 25th to the 75th percentile of the RV distribution in 2005 leads to an increase of almost 10% in green patents.

These results confirm the propositions outlined in the introduction, and are coherent with the conceptual framework of Section 2. The development of green technology relies on different types of regional know-how along the life cycle path. Specifically, unrelated variety exerts more influence at the beginning of the life cycle when technologies are still at early stages and knowledge recombination of cognitive distant knowledge is necessary for experimentation and trial and error. At early phases also R&D and human capital are fundamental to trigger patenting activity. However, in the maturity phase, when a dominant design is established, regional diversification is the main driver of green knowledge production though at a higher level of technological relatedness.

TABLE FIVE ABOUT HERE

5 Conclusions

The paper has explored empirically the relationship between local knowledge structures and the generation of environmental-related technology in the US over a thirty-year period. We framed the analysis in the life cycle heuristic to test whether the development of green technology benefits from specific types of agglomeration economies at different levels of technological relatedness. While prior literature in economic geography had

[18] In Appendix C we test the robustness of the assignment of green technologies to the life cycle stage. Using an alternative methodology to assign Env-Tech 2-digit technologies close to the mean value, the coefficient of RV in the diffusion phase is not significantly different from zero.

acknowledged the industry life cycle the present paper is, to the best of our knowledge, the first to employ such a heuristic to empirically study the spatial contingences of innovation processes.

Our empirical analysis yields two main findings. First, environment-related innovation is positively correlated with local knowledge base that is diversified across unrelated technological fields. This is coherent with the notion that green technology is on average more radical and complex than non-green technology, and that it draws upon knowledge inputs across cognitively distant domains (De Marchi, 2012; Barbieri et al., 2018; Vona et al, 2018). Second, we find that diversification across unrelated technological domains in local innovative activities favours green innovation mostly at early stages of development. Conversely, mature technologies benefit from diversification across related knowledge domains. This confirms our main conjecture and is consistent with Castaldi et al. (2015) concerning the role of local knowledge variety on technological innovation.

These results add to various scholarly and policy debates. First, we propose a novel empirical operationalisation to study environmental innovation within economic geography. Taking the cue from Truffer and Coenen (2012), we believe that the pressure of taking practical steps towards tackling climate change is upon regions and cities. Exposure to environmental phenomena however differs across territories, as does the ability to deal with the associated hazards (Perruchas et al, 2019). Add to the latter that green technologies are often treated as a homogenous block but, as our analysis shows, they differ substantially both from one another and from standard technologies (Barbieri et al., 2018; Quatraro and Scandura, 2019). While on the whole green technologies are at initial stages of development (OECD, 2011; Barbieri and Consoli, 2019), some are more mature than others (e.g. photovoltaics panels vs carbon capture). By reviving the notion that (green) technologies change over the life cycle as the selection environment alters their standing relative to knowledge domains and space, we shed light on the spatial diversity of invention capacity and of its evolution over time (Henning, 2019). In so doing, we provide a richer understanding of the strategies available (or not) to regions for tackling climate change by calling attention to the complexities at the interface of three sources of heterogeneity: environmental, spatial and technological.

Second, the articulation of technology maturity over degrees of relatedness validates empirically prior warnings on the perils of determinism (Boschma and Frenken, 2006). There is an incredibly rich range of knowledge configurations between the two extremes, related and unrelated, and there is scope for empirical research to delve into the connection between technology development and the organisation of local competences. We enter this debate by proposing an original method to capture the technology life cycle based on patent indicators. Although we focus on green technology, our method can be applied to any technology, and is scalable across levels of spatial aggregation (i.e. country, region, etc) as well as across technological domains (i.e. one-, two-, three-digits). Further, the paper provides initial insights into the dynamic links between technology and industry. While prior literature acknowledges that industry evolution is contingent on agglomeration economies, we find that also technology evolution depends on the regional knowledge structure. The present paper therefore paves the way for future research on another piece of the puzzle: do agglomeration economies affect industry dynamics through their impact on the knowledge generation process?

A third contribution is that we complement qualitative approaches rooted in the socio-technical transition agenda on sustainability. Given the common ground on evolutionary drivers of regional and industrial development, we analyse quantitative data through the lenses of a qualitative heuristic. In so doing, we cross-fertilise methods along the lines indicated by Truffer (2012) and Boschma (2017). What is integral to both approaches is the need to account for spatial contingencies that bring to bear on the capacity of cities, regions and countries to adapt production and consumption. This paper has identified a connection between the organisation of local knowledge and the differential state of development of green technology that advances the subfield of environmental economic geography (Tanner, 2014; Montresor and Quatraro, 2019).

Fourth, the method and the findings of the present paper are relevant for the policy debate on smart specialisation and on the design of strategies that assist regions in leveraging available assets to deal with their particular socio-economic challenges. As the European Commission acknowledges, successful smart specialisation calls for, among other things, bottom-up knowledge of the overall innovation support system.[19] The database created for the present study is a useful toolbox to aid the identification of local competitive strengths and of opportunities for entrepreneurial discovery.[20] To illustrate, targeting green technology that appears close to a region's prior specialisation may entail additional, unexpected, costs and trade-offs if the technology is at early stages of the life

[19] See i.e. https://s3platform.jrc.ec.europa.eu/what-is-smart-specialisation- (last accessed: 25 August, 2019)
[20] The data are freely available at https://www.greentechdatabase.com/ (last accessed: 16 December, 2019).

cycle and thus require high degrees of tacit know-how for successful implementation. Conversely, pursuing smart specialisation in mature domains may hinder entrepreneurial incentives if the developmental potential has been already widely exploited and the prospective returns are low, even if the technology related to the region's competence domain. All in all, there is much to be gained from the systematic articulation of technology life cycle and local knowledge structures.

Acknowledgements

We are indebted to the editor and four anonymous referees for constructive feedback. We are also grateful for comments received at the following events: SEEDS Workshop (Ferrara, 2018); GeoINNO Conference (Barcelona, 2018); IAERE Conference (Turin, 2018); INGENIO PhD Days (Valencia, 2018); EU-SPRI Conference (Paris, 2018); Technology Transfer Conference (Valencia, 2018); Spanish Regional Sciences Association (Valencia, 2018); IAERE Conference (Udine, 2019); Rethinking Clusters Workshop (Padua, 2019); EMAEE Conference (Brighton, 2019). DC acknowledges the financial support of Consejo Superior de Investigaciones Científicas (CSIC) "Ayuda de Incorporación para Científicos Titulares" (OEP2014) and of the Ministerio de Economia Industria y Competitividad "Programa Estatal de I+D+i Orientada a los Retos de la Sociedad, 2017" (ECO2017- 86976R). FP acknowledges the financial support of European H2020 Project RISIS2 (grant agreement nº 824091). All errors and omissions are our own.

References

Abernathy, W.J., Utterback, J. (1978) Patterns of industrial innovation. *Technology Review* 50, pp. 41-47

Abernathy, W. J., and Clark, K. B. (1985). Innovation: Mapping the winds of creative destruction. *Research Policy*, *14*(1), pp. 3-22.

Achilladelis, B., Schwarzkopf, A., and Cines, M. (1990). The dynamics of technological innovation: The case of the chemical industry. *Research Policy*, *19*(1), pp. 1-34.

Achilladelis, B. (1993). The dynamics of technological innovation: The sector of antibacterial medicines. Research Policy, 22(4), 279–308.

Adam B., J. (1986). *Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value*. Cambridge, MA: National Bureau of Economic Research.

Agarwal, R., and Gort, M. (2002). Firm and Product Life Cycles and Firm Survival. *American Economic Review*, *92*(2), pp. 184-190.

Andersen, B. (1999). The hunt for S -shaped growth paths in technological innovation: a patent study. *Journal of Evolutionary Economics*, *9*(4), pp. 487-526.

Antonioli, D., Borghesi, S., & Mazzanti, M. (2016). Are regional systems greening the economy? Local spillovers, green innovations and firms' economic performances. *Economics of Innovation and New Technology*, 25(7), 692-713.

Arthur, W.B. (2007). The structure of invention. *Research Policy*, *36*(2), pp. 274-287.

Attaran, M. (1986). Industrial diversity and economic performance in U.S. areas. *The Annals of Regional Science*, *20*(2), pp. 44-54.

Audretsch, D. B., and Feldman, M. P. (1996). Innovative clusters and the industry life cycle. *Review of Industrial Organization*, *11*(2), pp. 253-273.

Balland, P. A., and Rigby, D. (2017). The Geography of Complex Knowledge. *Economic Geography*, *93*(1), pp. 1-23.

Barbieri, N., Ghisetti, C., Gilli, M., Marin, G., & 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. (2018). Knowledge sources and impacts on subsequent inventions: Do green technologies differ from non-green ones? *SPRU Working Paper Series 2018-11* (http://sci-hub.tw/10.2139/ssrn.3164197)

Barbieri, N., and Consoli, D. (2019). Regional diversification and green employment in US Metropolitan Areas. *Research Policy*, *48*(3), 693-705.

Basalla, G. (1989). *The Evolution of Technology.* Cambridge University Press.

Bishop, P., and Gripaios, P. (2010). Spatial Externalities, Relatedness and Sector Employment Growth in Great Britain. *Regional Studies*, *44*(4), pp. 443-454.

Boschma, R. A., and Frenken, K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography, 6*(3), pp. 273-302.

Boschma, R., and Iammarino, S. (2009). Related Variety, Trade Linkages, and Regional Growth in Italy. *Economic Geography, 85*(3), pp. 289-311.

Boschma, R., Coenen, L., Frenken, K., and Truffer, B. (2017). Towards a theory of regional diversification: combining insights from Evolutionary Economic Geography and Transition Studies. *Regional Studies, 51*(1), pp. 31-45.

Brunnermeier, S. B., & Cohen, M. A. (2003). Determinants of environmental innovation in US manufacturing industries. *Journal of Environmental Economics and Management, 45*(2), pp. 278–293.

Cainelli, G. and Iacobucci, D. (2012), Agglomeration, Related Variety, and Vertical Integration. *Economic Geography*, 88, pp. 255-277.

Cainelli, G., De Marchi, V., & Grandinetti, R. (2015). Does the development of environmental innovation require different resources? Evidence from Spanish manufacturing firms. *Journal of Cleaner Production, 94*, 211-220.

Callon, M. (1998). An Essay on Framing and Overflowing: Economic Externalities Revisited by Sociology. *The Sociological Review, 46*(1_suppl), pp. 244-269.

Castaldi, C., Frenken, K., and Los, B. (2015). Related Variety, Unrelated Variety and Technological Breakthroughs: An analysis of US State-Level Patenting. *Regional Studies*, *49*(5), pp. 767-781.

Chang, S.-H., and Fan, C.-Y. (2016). Identification of the technology life cycle of telematics: A patent-based analytical perspective. *Technological Forecasting and Social Change, 105*, pp. 1-10.

Chang, C. P., Wen, J., Zheng, M., Dong, M., & Hao, Y. (2018). Is higher government efficiency conducive to improving energy use efficiency? Evidence from OECD countries. *Economic Modelling*, 72, pp. 65-77.

Charlot, S., Crescenzi, R., and Musolesi, A. (2015). Econometric modelling of the regional knowledge production function in Europe. *Journal of Economic Geography, 15*(6), pp. 1227-1259.

Christensen, Clayton M (1997). Innovator's Dilemma:When new technologies cause great firms to fail. *Harvard Business School Press Books.*

Constant (1980). The origins of the turbojet revolution. *Johns Hopkins studies in the history of technology.*

Content, J., and Frenken, K. (2016). Related variety and economic development: a literature review. *European Planning Studies, 24*(12), pp. 2097-2112.

Crescenzi, R., Rodriguez-Pose, A., and Storper, M. (2007). The territorial dynamics of innovation: a Europe United States comparative analysis. *Journal of Economic Geography, 7*(6), pp. 673-709.

Crescenzi, R., and Rodríguez-Pose, A. (2012). Infrastructure and regional growth in the European Union*. *Papers in Regional Science, 91*(3), pp. 487-513.

Crescenzi, R., Gagliardi, L., and Iammarino, S. (2015). Foreign multinationals and domestic innovation: Intra-industry effects and firm heterogeneity. *Research Policy, 44*(3), pp. 596-609.

De Marchi, V. (2012). Environmental innovation and R&D cooperation: Empirical evidence from Spanish manufacturing firms. *Research Policy, 41*(3), pp. 614-623.

Desrochers, P., and Leppälä, S. (2011). Creative Cities and Regions: The Case for Local Economic Diversity. *Creativity and Innovation Management, 20*(1), pp. 59-69.

Driscoll, J. C., and Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics, 80*(4), pp. 549-560.

Duranton, G., and Puga, D. (2001). Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. *American Economic Review, 91*(5), pp. 1454-1477.

Essletzbichler, J. (2007). Diversity, Stability and Regional Growth in the United States, 1975 - 2002. In K. Frenken (Ed.). *Applied Evolutionary Economics and Economic Geography*, Edward Elgar Publishing , Cheltenham, UK.

Feldman, M. P., and Audretsch, D. B. (1999). Innovation in cities:: Science-based diversity, specialization and localized competition. *European Economic Review, 43*(2), pp. 409-429.

Feldman, M. P. and A. Graddy Reed. (2014) Local champions: entrepreneurs' transition to philanthropy and the vibrancy of place. In: M.L. Taylor, R.J. Strom and D. O. Renz

(Eds.) *Handbook of Research on Entrepreneurs' Engagement in Philanthropy*, Edward Elgar, Cheltenham UK, Northampton MA, US, pp. 43-71.

Fleming, L. (2001). Recombinant Uncertainty in Technological Search. *Management Science, 47*(1), pp. 117-132.

Foster, R. N. (1986). *Innovation : the attacker's advantage.* Simon & Schuster.

Fredriksson, P. G., & Vollebergh, H. R. J. (2009). Corruption, federalism, and policy formation in the OECD: the case of energy policy. *Public Choice, 140*(1), pp. 205–221.

Frenken, K., and Boschma, R. A. (2007). A theoretical framework for evolutionary economic geography: industrial dynamics and urban growth as a branching process. *Journal of economic geography, 7*(5), pp. 635-649.

Gao, L., Porter, A. L., Wang, J., Fang, S., Zhang, X., Ma, T., . . . Huang, L. (2013). Technology life cycle analysis method based on patent documents. *Technological Forecasting and Social Change, 80*(3), pp. 398-407.

Ghisetti, C. and Pontoni, F. (2015), Investigating policy and R&D effects on environmental innovation: A meta-analysis, *Ecological Economics, 118*(C), pp. 57-66.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy, 100*(6), pp. 1126-1152.

Gort, M., and Klepper, S. (1982). Time Paths in the Diffusion of Product Innovations. *The Economic Journal, 92*(367), p. 630.

Griliches, Z. (1979). Issues in Assessing the Contribution of Research and Development to Productivity Growth. *The Bell Journal of Economics, 10*(1), p. 92.

Griliches, Zvi, 1990. Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, *28*(4), p. 1661.

Hall, B. H., and Helmers, C. (2013). Innovation and diffusion of clean/green technology: Can patent commons help? *Journal of Environmental Economics and Management, 66*(1), pp. 33-51.

Hall, B., and Trajtenberg, M. (2004). *Uncovering GPTS with Patent Data.* Cambridge, MA: National Bureau of Economic Research.

Harrison, B., Kelley, M., & Gant, J. (1996). Specialization Versus Diversity in Local Economies: The Implications for Innovative Private-Sector Behavior. *Cityscape, 2*(2), 61-93.

Hascic, I., & Migotto, M. (2015). Measuring environmental innovation using patent data. *OECD Environment Working Papers* 89.

Haupt, R., Kloyer, M., and Lange, M. (2007). Patent indicators for the technology life cycle development. *Research Policy, 36*(3), pp. 387-398.

Henderson, R. (1995). Of life cycles real and imaginary: The unexpectedly long old age of optical lithography. *Research Policy, 24*(4), pp. 631-643.

Henderson, V., Kuncoro, A., and Turner, M. (1995). Industrial Development in Cities. *Journal of Political Economy, 103*(5), pp. 1067-1090.

Henning, M. (2019). Time should tell (more): evolutionary economic geography and the challenge of history. *Regional Studies, 53*(4), 602-613.

Jacobs, J. (1969). *The economy of cities.* Vintage Books.

Jacquemin, A. P., and Berry, C. H. (1979). Entropy Measure of Diversification and Corporate Growth. *Journal of Industrial Economics, 27*(4), pp. 359-69.

Jaffe, A. B. (1986). Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value. *The American Economic Review, 76*(5), pp. 984-1001.

Jaffe, A. B. (1989). Real Effects of Academic Research. *The American Economic Review, 79*(5), pp. 957-970.

Joskow P. L. & Kohn E. (2002). A Quantitative Analysis of Pricing Behavior in California's Wholesale Electricity Market During Summer 2000, *The Energy Journal, 0*(4), pp. 1-35.

Kelley, M. R., and Helper, S. (1999). Firm Size And Capabilities, Regional Agglomeration, And The Adoption Of New Technology. *Economics of Innovation and New Technology, 8*(1-2), pp. 79-103.

Klepper, S. (1997). Industry Life Cycles. *Industrial and Corporate Change, 6*(1), pp. 145-181.

Klepper, S. (1996). Entry, Exit, Growth, and Innovation over the Product Life Cycle. *The American Economic Review, 86*(3), pp. 562-583.

Hughes, T. P. (1983). *Networks of Power: Electrification in Western Society, 1880-1930*. John Hopkins University Press, Baltimore and London.

Laplante, B., & Rilstone, P. (1996). Environmental Inspections and Emissions of the Pulp and Paper Industry in Quebec. *Journal of Environmental Economics and Management, 31*(1), pp. 19–36.

Lee, C., Cho, Y., Seol, H., and Park, Y. (2012). A stochastic patent citation analysis approach to assessing future technological impacts. *Technological Forecasting and Social Change, 79*(1), pp. 16-29.

Lee, C., Kim, J., Kwon, O., and Woo, H.-G. (2016). Stochastic technology life cycle analysis using multiple patent indicators. *Technological Forecasting and Social Change, 106*, pp. 53-64.

Lim, M., Zhou, Y., Wang, L., Rudolph, V., and Lu, G. Q. (2009). Development and potential of new generation photocatalytic systems for air pollution abatement: an overview. *Asia-Pacific Journal of Chemical Engineering, 4*(4), pp. 387-402.

March, J. G. and Simon, Herbert A., (1958). Organizations. University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship.

Markusen, J. R., and Svensson, L. E. (1985). Trade in Goods and Factors with International Differences in Technology. *International Economic Review*, *26*(1), p. 175.

Marshall, A. (1920). Industry and Trade. *Journal of the Royal Statistical Society*.

Mazzanti, M., Marin, G., Mancinelli, S., & Nicolli, F. (2015). Carbon dioxide reducing environmental innovations, sector upstream/downstream integration and policy: evidence from the EU. *Empirica*, 42(4), 709-735.

Montresor, S. and Quatraro, F. (2017). Regional branching and Key enabling technologies: Evidence from European patent data. *Economic Geography, 93(4),* pp. 367-396.

Montresor, S., and Quatraro, F. (2019). Green technologies and Smart Specialisation Strategies: a European patent-based analysis of the intertwining of technological relatedness and key enabling technologies. *Regional Studies (forthcoming).*

Neffke, F., Henning, M., Boschma, R., Lundquist, K.-J., and Olander, L.-O. (2011a). The Dynamics of Agglomeration Externalities along the Life Cycle of Industries. *Regional Studies, 45*(1), pp. 49-65.

Neffke, F., Henning, M., and Boschma, R. (2011b). How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions. *Economic Geography, 87*(3), pp. 237-265.

Nelson, R. R., and Winter, S. G. (1982). *An evolutionary theory of economic change.* Belknap Press of Harvard University Press.

Nooteboom, B. (2002). A balanced theory of sourcing, collaboration and networks. *ERIM Report Series Research in Management*.

Norton, R. D., and Rees, J. (1979). The product cycle and the spatial decentralization of American manufacturing. *Regional Studies, 13*(2), pp. 141-151.

OECD. (2011). *Towards Green Growth*. OECD.

OECD. (2015). *Monitoring the transition to a low-carbon economy: a strategic approach to local development.* OECD.

Perruchas, F., Consoli, D. and Barbieri, N. (2019) Specialisation, Diversification and the Ladder of Green Technology Development. *SPRU Working Paper Series 2019-07*.

Ponds, R., Oort, F. v., and Frenken, K. (2010). Innovation, spillovers and university-industry collaboration: an extended knowledge production function approach. *Journal of Economic Geography, 10*(2), pp. 231-255.

Popp, D. (2005). Lessons from patents: Using patents to measure technological change in environmental models. *Ecological Economics*, 54(2–3), 209–226.

Popp, D., Newell, R. G., & Jaffe, A. B. (2010). Energy, the Environment, and Technological Change. In N. Rosenberg & B. Halland (Eds.), Handbook of the Economics of Innovation- Vol-II (pp. 873–938). Burlington, USA: Academic Press.

Quatraro, F., and Scandura, A. (2019). Academic inventors and the antecedents of green technologies. A regional analysis of Italian patent data. *Ecological Economics*, *156*, pp. 247-263.

Rigby, D. L., and Essletzbichler, J. (1997). Evolution, process variety, and regional trajectories of technological change in US manufacturing. *Economic Geography*, *73*(3), pp. 269-284.

Rosenberg, N. (1976). *Perspectives on technology.* Cambridge University Press.

Schumpeter, J. A. (1939). *Business cycles : a theoretical, historical, and statistical analysis of the capitalist process.* Martino Pub.

Storper M., and Walker R. (1989). *The Capitalist Imperative. Territory, Technology and Industrial Growth*. Basil Blackwell, New York, NY.

Tanner, A. N. (2014). Regional branching reconsidered: Emergence of the fuel cell industry in European regions. Economic Geography, 90(4), 403–427.

The Boston Consulting Group (2015). *Steel as a Model for Sustainable Metal Industry in 2050.*

Theil, H. (1972). *Statistical decomposition analysis. With applications in the social and administrative sciences.* Amsterdam: North-Holland Pub. Co.

Truffer, B., and Coenen, L. (2012). Environmental Innovation and Sustainability Transitions in Regional Studies. *Regional Studies, 46*(1), pp. 1-21.

US Department of Energy (2017). *Fuel Cell Technologies Market Report 2016.* Washington, DC.

US National Academy of Sciences, National Academy of Engineering, and National Research Council (2010) *Electricity from Renewable Resources: Status, Prospects and Impediments*. Washington, DC: The National Academies Press.

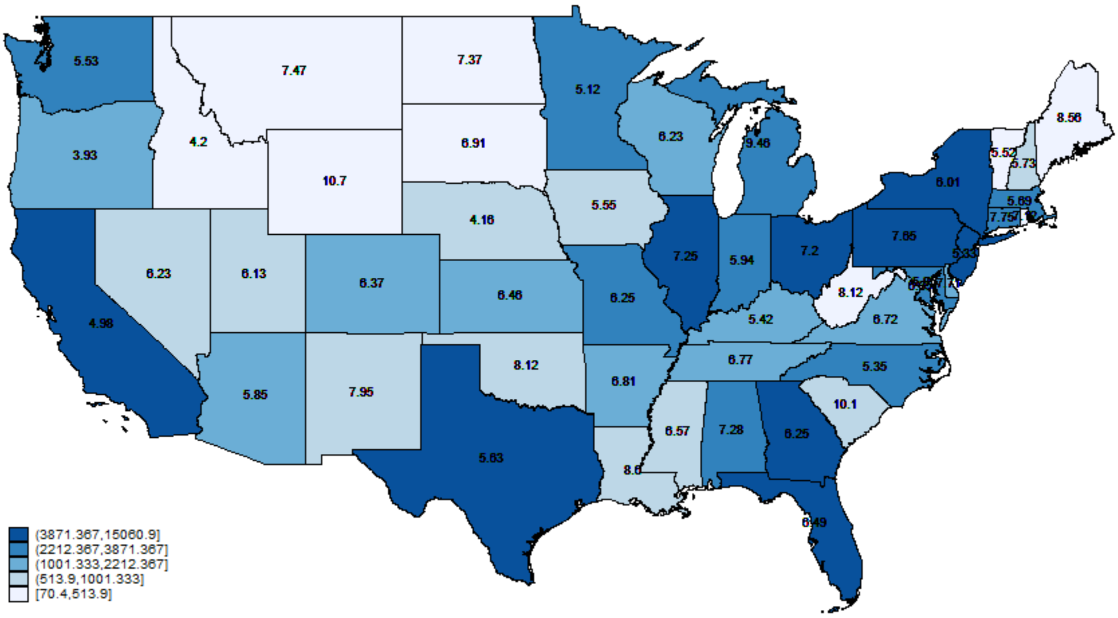
Vona, F. and Consoli, D. (2015). Innovation and skill dynamics: a life-cycle approach. *Industrial and Corporate Change*, 24(6):1393-1415.

Vona, F., Marin, G., Consoli, D., and Popp, D. (2018). Environmental regulation and green skills: an empirical exploration. *Journal of the Association of Environmental and Resource Economists*, *5*(4): 713-753.

Weitzman, M. L. (1998). Recombinant Growth. *The Quarterly Journal of Economics*, *113*(2), pp. 331-360.

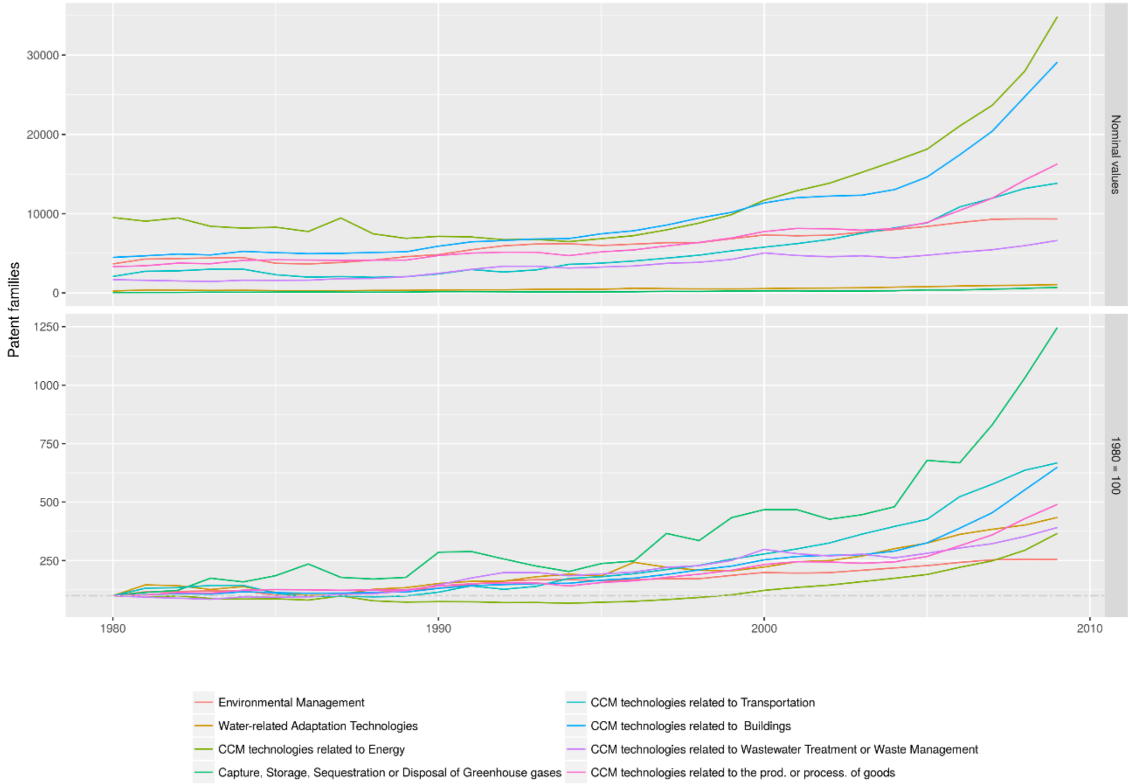
Figures

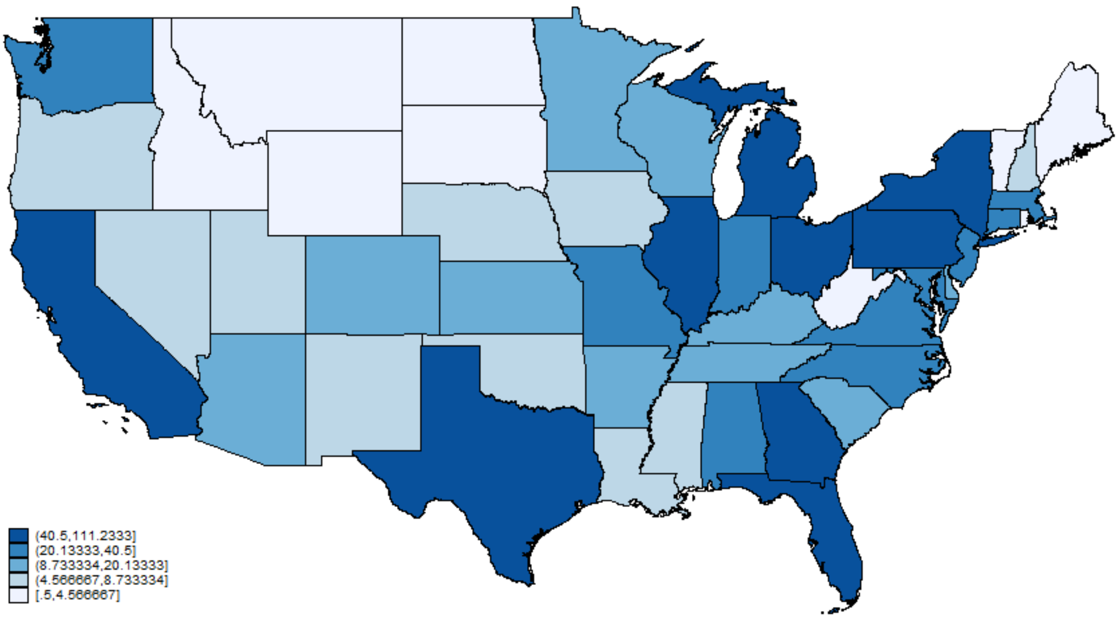
Figure 1. Quintiles of total patent families and percentage of green patents (average 1980-2010)



Notes: Darker colours correspond to top quintiles of the distribution of total patents. 48 US federal states and District of Columbia are included in the maps. Alaska and Hawaii are. The numbers within US states borders correspond to the percentage of green patent families. The cartographic boundary shapefile is provided by the US Census Bureau (Accessed in 2018). Source: Own elaboration

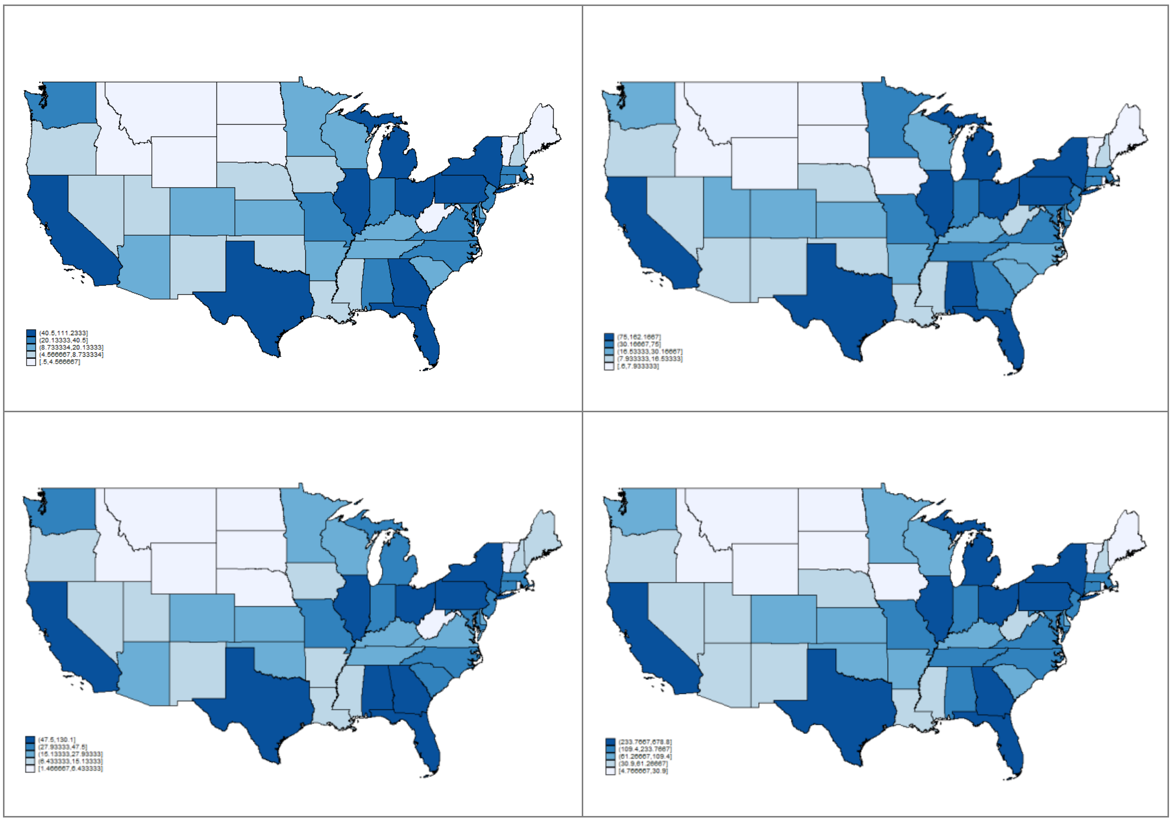
Figure 2. Evolution of the number of green patent families by Env-Tech families, 1980 – 2009. Top panel: nominal values; bottom panel: 1980= 100.





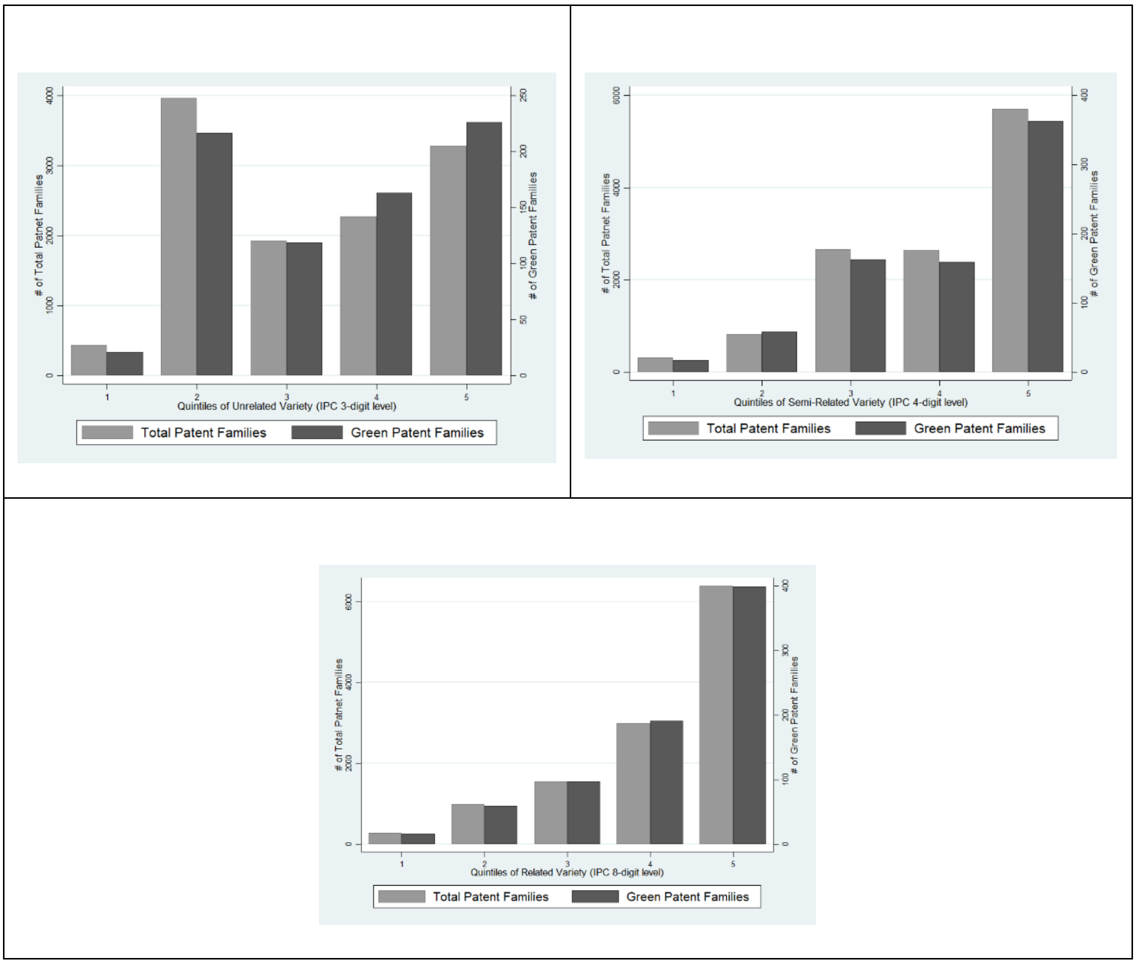
Source: Own elaboration

Figure 3. Quintiles of green patent families over technology life cycle stages (average 1980-2009)



Note: Darker colours correspond to top quintiles. 48 US federal states and District of Columbia are included in the maps. Alaska and Hawaii are not included. The cartographic boundary shapefile is provided by the US Census Bureau (Accessed in 2018). Source: Own elaboration.

Figure 4. Distribution of green and total patent families over quintiles of Unrelated, Semi-Related and Related variety (average 1980-2009)



Tables

Table 1. Life cycle stages

		Ubiquity	
Patenting intensity		*Low*	*High*
	Low	**Emergence (1)**	**Diffusion (3)**
	High	**Development (2)**	**Maturity (4)**

Notes: The numbers in parenthesis are labels for the stage of development. The same notation is used in Table 2.

Table 2. Life cycle stages and number of patent families by Env-Tech 2-digit codes

ID	**ENV-TECH**	**1980**	**1990**	**2000**	**2010**	**# of patent families 1980-2009**
1.1	AIR POLLUTION ABATEMENT	4	4	4	4	97,729
1.2	WATER POLLUTION ABATEMENT	3	4	4	4	55,246
1.3.	WASTE MANAGEMENT	3	3	4	4	44,353
1.4	SOIL REMEDIATION	1	1	3	3	3,813
1.5	ENVIRONMENTAL MONITORING	1	1	1	1	3,523
2.1	DEMAND-SIDE TECH (water conservation)	1	3	3	3	13,507
2.2	SUPPLY-SIDE TECH (water availability)	1	1	1	3	3,716
4.1	RENEWABLE ENERGY GENERATION	4	4	4	4	158,579
4.2	ENERGY GENERATION FROM FUELS OF NON-FOSSIL ORIGIN	1	3	3	4	24,354
4.3	COMBUSTION TECH WITH MITIGATION POTENTIAL	1	1	1	3	11,519
4.4	NUCLEAR ENERGY	2	2	1	1	33,734
4.5	EFFICIENCY IN ELECTRICAL POWER GENERATION, TRANSMISSION OR DISTRIBUTION	1	2	1	1	18,635
4.6	ENABLING TECH IN ENERGY SECTOR	1	2	2	2	152,431
4.7	OTHER ENERGY CONVERSION OR MANAGEMENT SYSTEMS REDUCING GHG EMISSIONS	1	1	1	3	2,788
5.1	CO2 CAPTURE OR STORAGE (CCS)	1	1	1	3	4,930
5.2	CAPTURE OR DISPOSAL OF GREENHOUSE GASES OTHER THAN CARBON DIOXIDE (N2O, CH4, PFC, HFC, SF6)	1	1	1	3	2,353
6.1	ROAD TRANSPORT	2	4	2	2	147,386
6.2	RAIL TRANSPORT	1	1	1	1	2,568
6.3	AIR TRANSPORT	1	1	1	3	12,984
6.4	MARITIME OR WATERWAYS TRANSPORT	1	1	1	3	3,917
6.5	ENABLING TECH IN TRANSPORT	1	1	1	2	10,678
7.1	INTEGRATION OF RENEWABLE ENERGY SOURCES IN BUILDINGS	1	1	1	4	20,883
7.2	ENERGY EFFICIENCY IN BUILDINGS	1	3	4	4	91,836
7.3	ARCHITECTURAL OR CONSTRUCTIONAL ELEMENTS IMPROVING THE THERMAL PERFORMANCE OF BUILDINGS	1	1	1	1	2,879
7.4	ENABLING TECH IN BUILDINGS	4	4	4	4	214,003
8.1	WASTEWATER TREATMENT	1	3	4	4	40,568
8.2	SOLID WASTE MANAGEMENT	3	3	4	4	66,033
8.3	ENABLING TECH OR TECH WITH A POTENTIAL OR INDIRECT CONTRIBUTION TO GHG MITIGATION	1	1	1	1	3,499
9.1	TECH RELATED TO METAL PROCESSING	3	3	3	4	39,059
9.2	TECH RELATING TO CHEMICAL INDUSTRY	1	4	4	4	67,329
9.3	TECH RELATING TO OIL REFINING AND PETROCHEMICAL INDUSTRY	1	1	1	3	4,149
9.4	TECH RELATING TO THE PROCESSING OF MINERALS	1	3	1	3	12,215
9.5	TECH RELATING TO AGRICULTURE, LIVESTOCK OR AGROALIMENTARY INDUSTRIES	1	3	1	3	8,980
9.6	TECH IN THE PRODUCTION PROCESS FOR FINAL INDUSTRIAL OR CONSUMER PRODUCTS	1	1	2	4	58,171
9.7	CLIMATE CHANGE MITIGATION TECH FOR SECTOR-WIDE APPLICATIONS	1	1	1	1	2,249
9.8	ENABLING TECH WITH A POTENTIAL CONTRIBUTION TO GHG EMISSIONS MITIGATION	1	1	1	4	24,302

Notes: ID and ENV-TECH correspond to green technology groups listed in OECD (2016). Numbers in the columns indicate the life cycle stage of green technologies: 1="Emergence", 2="Development", 3="Diffusion", 4="Maturity" (as per Table 1). Dark colours are associated to later stages of the technology life cycle.

Table 3. Descriptive statistics

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
UV	*Unrelated variety at 3-digit level*	1,470	3.773	.221	2.832	4.204
SRV	*Semi-Related Variety at 4-digit level*	1,470	1.248	.205	.268	1.528
RV	*Related Variety at 8-digit level*	1,470	1.453	.361	.246	1.916
GP	*Number of Green patent families*	1,470	27.69	26.64	0	300.94
Tot Pat	*Number of Total patent families*	1,470	429.6	351.2 4	36.18	2810.15
Emergence	*Number of Green patent families, Emergence stage*	1,470	4.451	4.781	0	51.61
Development	*Number of Green patent families, Development stage*	1,470	6.821	9.026	0	95.21
Diffusion	*Number of Green patent families, Diffusion stage*	1,470	6.163	6.345	0	83.79
Maturity	*Number of Green patent families, Maturity stage*	1,470	24.08	25.93	0	320.18
R&D	*Research and Development expenditures (w.r.t. GDP)*	1,470	.014	.011	.001	.066
HC	*% Population with bachelor degree or more*	1,470	.057	.021	.0321	.541
R&D Neighb	*Research and Development expenditures in neighbouring states (w.r.t. GDP)*	1,470	.015	.007	.002	.047
HC Neighb	*% Population with bachelor degree or more in neighbouring states*	1,470	.055	.007	.037	.093
Pop Dens	*Population Density*	1,470	4.80	1.476	1.53	9.14
Inspections	*Number of inspections over the number of establishments*	1,470	.22	0.33	0	1.98
Energy intensity	*Energy consumption over GDP*	1,470	15.41	8.43	1.91	61.15

Notes: Number of States: 49; Coverage: 1980-2009

Table 4. Regression results

	(1)	(2)	(3)	(4)	(5)	(6)
	GP (log)	Non-GP (log)	GP (log)	Non-GP (log)	GP (log)	Non-GP (log)
UV (log)	1.339***	-1.093**	1.330***	-1.098**	1.541***	-1.018**
	(0.455)	(0.528)	(0.443)	(0.513)	(0.469)	(0.491)
SRV (log)	0.289	0.218***	0.277	0.205***	0.295	0.205***
	(0.179)	(0.0690)	(0.179)	(0.0716)	(0.186)	(0.0729)
RV (log)	0.387**	0.523***	0.384**	0.521***	0.414**	0.535***
	(0.146)	(0.0943)	(0.145)	(0.0946)	(0.168)	(0.103)
R&D (log)			0.0222	0.00674	0.0242	0.00650
			(0.0188)	(0.00971)	(0.0156)	(0.00903)
HC (log)			0.108*	0.113*	0.0668	0.0754**
			(0.0565)	(0.0602)	(0.0562)	(0.0365)
R&D Neighb (log)					0.0317	0.0196
					(0.0617)	(0.0249)
HC Neighb (log)					0.285	0.208**
					(0.210)	(0.0791)
Pop Dens					0.000868**	0.000421***
					(0.000370)	(0.000113)
Inspections					0.0844	
					(0.0500)	
Energy Intensity					-0.00326	
					(0.00627)	
State FE	Y	Y	Y	Y	Y	Y
Time Dummies	Y	Y	Y	Y	Y	Y
Random growth	Y	Y	Y	Y	Y	Y
Obs.	1466	1470	1466	1470	1466	1470
R2	0.889	0.974	0.890	0.975	0.892	0.975
F	1416331.9	127000.0	2445371.2	197715.2	2115086.2	15295553.0

*Notes: The analysis covers 48 US Federal States and the District of Columbia over 1980-2009. Driscoll and Kraay's (1998) standard errors, robust to heteroskedasticity and serial and spatial correlation, in parentheses. * p < 0.1; **p < 0.05; *** p < 0.01.*

Table 5. Regression results over the life cycle

	GP (log)	Emergence	Development	Diffusion	Maturity
UV (log)	1.541***	1.133**	1.380**	0.753	0.755
	(0.469)	(0.446)	(0.649)	(0.804)	(0.464)
SRV (log)	0.295	-0.395	0.812***	0.148	-0.254
	(0.186)	(0.291)	(0.189)	(0.229)	(0.150)
RV (log)	0.414**	0.495	0.545***	0.458*	0.587***
	(0.168)	(0.311)	(0.163)	(0.253)	(0.0888)
R&D (log)	0.0242	0.0815***	0.0266	0.0588**	-0.0311
	(0.0156)	(0.0273)	(0.0440)	(0.0221)	(0.0212)
HC (log)	0.0668	-0.147*	0.305**	0.153*	-0.0331
	(0.0562)	(0.0829)	(0.137)	(0.0772)	(0.0517)
R&D Neighb (log)	0.0317	0.153***	0.0664	0.115	-0.0679
	(0.0617)	(0.0417)	(0.103)	(0.0821)	(0.0437)
HC Neighb (log)	0.285	-0.209	0.263	-0.0213	0.792***
	(0.210)	(0.304)	(0.898)	(0.289)	(0.270)
Pop Dens	0.000868**	0.00196***	0.000925***	0.00130***	0.000699
	(0.000370)	(0.000244)	(0.000327)	(0.000353)	(0.000435)
Inspections	0.0844	-0.0485	0.180**	-0.112*	0.0591
	(0.0500)	(0.0786)	(0.0800)	(0.0615)	(0.0496)
Energy Intensity	-0.00326	0.00930	-0.0154	0.000977	0.00677
	(0.00627)	(0.0124)	(0.0132)	(0.00626)	(0.00412)
State FE	Y	Y	Y	Y	Y
Time Dummies	Y	Y	Y	Y	Y
Random growth	Y	Y	Y	Y	Y
Obs.	1466	1392	1371	1424	1452
R2	0.892	0.591	0.794	0.687	0.906
F	2115086.2	128218.5	238252.4	528859.7	634091.9

*Notes: The analysis covers 48 US Federal States and the District of Columbia over 1980-2009. Driscoll and Kraay's (1998) standard errors, robust to heteroskedasticity and serial and spatial correlation, in parentheses. * p < 0.1; **p < 0.05; *** p < 0.01.*

APPENDIX A (Online publication) – Missing inventor's address

Before geo-localisation we collect all the inventors' addresses from the PATSTAT database. Two main issues arise. First, although the European Patent Office (EPO) assigns an unambiguous ID to each applicant or inventor, we may still find multiple IDs for the same person due to misspelling, name variations, second names, etc. For instance, the inventor's name may appear as John Paul Smith, J. Smith or J.P. Smith and be assigned to different patents. Second, address information is provided in PATSTAT only for some inventors. In relation to the first issue, address information for an inventor may be provided for some IDs and missing for others. For example, address information may be provided for John Paul Smith and not for J. Smith due to differences in their IDs.

To reduce the number of inventors/applicants with a missing address we exploit the information on the patent family – our unit of analysis. Within each patent family we link multiple inventors' IDs assuming that they are the same person based on a string-matching indicator. We calculate the Levenshtein distance between the inventor name for which the address is provided and all the other names with missing information within the patent family. We consider two or more inventors as the same person if the indicator is below three. This means that their full names differ for less than three characters. Then, if the address information is provided for one of these inventors we assign it also to the other IDs for which this information is not provided (even though they have different IDs). For example, the same patent family can feature two inventors with different IDs, the first with a complete address, the second with a missing one: "Gehri, Martin Christian Adrian" and "GEHRI, MARTIN, CHRISTIAN, ADRIAN". As the Levenshtein distance between the two names is less than 3 when both strings are converted to uppercase, we assume it is the same person and we use the complete address to fill the missing one.

APPENDIX B (Online publication) – Technology Life Cycle indicator

Prior literature offers insights into methodologies to assess the stage of development of technologies through patent data. Haupt et al. (2007) use patent indicators and empirically test their difference along the technology life cycle stages. Although they do not directly use patent indicators to detect the stage of development of technologies, they show that these indicators follow specific patterns depending on technology development over life cycle stages that are defined a priori by a pool of experts and literature review. Other studies directly employ patent indicators (Gao et al., 2013; Chang and Fan, 2016). These works define life cycle stages of a benchmark technology through expert interviews and assess the trends of patent indicators over its technological evolution. Subsequently, these patent indicators are compared with those of the benchmark technology and the life cycle stage of the latter is then assigned to the former. Finally, stochastic techniques have been used to measure technology life cycle. Lee et al. (2012; 2016) run Hidden Markov Models to analyse patent indicators over time. Such a technique allows calculating the highest probability path that yields the most probable stage of development at each step.

Our focuses on a broad number of different environmental-related technologies reduces the chances to identify benchmark technologies – even with the contribution of a pool of experts. Moreover, in our opinion the stage of technology development should account for spatial diffusion, and not just for intensity of patenting. Finally, we aim at developing a technology life cycle indicator that provides values for broad technological fields not just single inventions.

As described in Section 3.4 we develop our measure of technology life cycle based on two indicators, i.e. the geographical ubiquity and patenting intensity. We calculate these indicators using worldwide patent families for each macro-technology reported in the Env-Tech classification. Figure A1 shows the life cycle of green technologies over the entire period of the analysis (1980-2009). We can observe that the indicator captures the heterogeneity that characterises green technologies allowing for non-linear transition between life cycle stages. For instance, ENV-TECH 7.1 “Integration of renewable energy sources in buildings” falls in the emergence stage until 2000 moving to the diffusing phase until maturity is reached in 2008. Green technologies aimed at reducing the environmental impact of nuclear energy follow an opposite pattern starting in the development phase moving to the emergence stage from 1990 onwards.

Some illustrative examples are provided in Figure A2. Technologies related to renewable energy generation exhibit a stable level of patenting activity since the period 1980-1989, while geographical ubiquity reaches the highest value among all the other technologies. This is in line with what we expect from a set of technologies in a diffusion, or mature, stage (National Academy of Sciences et al, 2010). On the other hand, a small number of countries contribute to the enabling technologies in transport (application of fuel cell or hydrogen technology to transportation and charging of electric vehicle) but patenting activity is increasing over time, meaning that these technologies are not mature but still in a development phase, in line with the evidence available (i.e. US Department of Energy, 2010). The other three technologies (air pollution abatement – 1.1, CO2 capture and storage – 5.1 and technologies related to metal processing – 9.1) in Figure A2 are instances of a shift from development towards maturity in that they exhibit sustained growth in patenting during the whole period while geographical ubiquity increases only over the last two decades, and in line with prior empirical studies (Lim et al., 2009). This pattern differs from that of technologies related to efficiency and reduction of greenhouse gas emissions in metal processing (9.1): between 1980-1989 and 1990-1999 patenting is stable and spread over a higher number of countries, while in the last decade, ubiquity diminishes and patenting activity grows again. This trajectory suggests a future change in the trend of the life cycle of these technologies (The Boston Consulting Group, 2015).

All the technologies follow a similar path, but some are more advanced in the TLC than others. For example, even if air pollution abatement and CO2 capture or storage are moving toward the diffusion stage, their movements start later compared to the average of all the other technologies. To characterize this evolution in the broader context of all green technologies, we calculate the average value of ubiquity and patenting growth rate for all the GT in each time period. The combination of these two characteristics gives rise to four different regimes (Table 1). "Emergence" technologies have patenting intensity and ubiquity below average; "development" technologies exhibit above average patenting and below average ubiquity; technologies in "diffusion" are above average in both intensity and ubiquity; in the "maturity" ubiquity is above average and patenting below the average of all the technologies in the same period. Figure A3 illustrates the 4 phases of TLC during the period 2000-2009 for the technologies shown in Figure A1 (dashed lines indicate mean values). In this example, CO2 capture or storage (5.1) and enabling technologies in transport (6.5) in the "emergence" phase, air pollution abatement (1.1) in the "development" phase, renewable energy generation (4.1) would be in the "diffusion" phase and technologies related to metal processing (9.1) in the "maturity" phase.

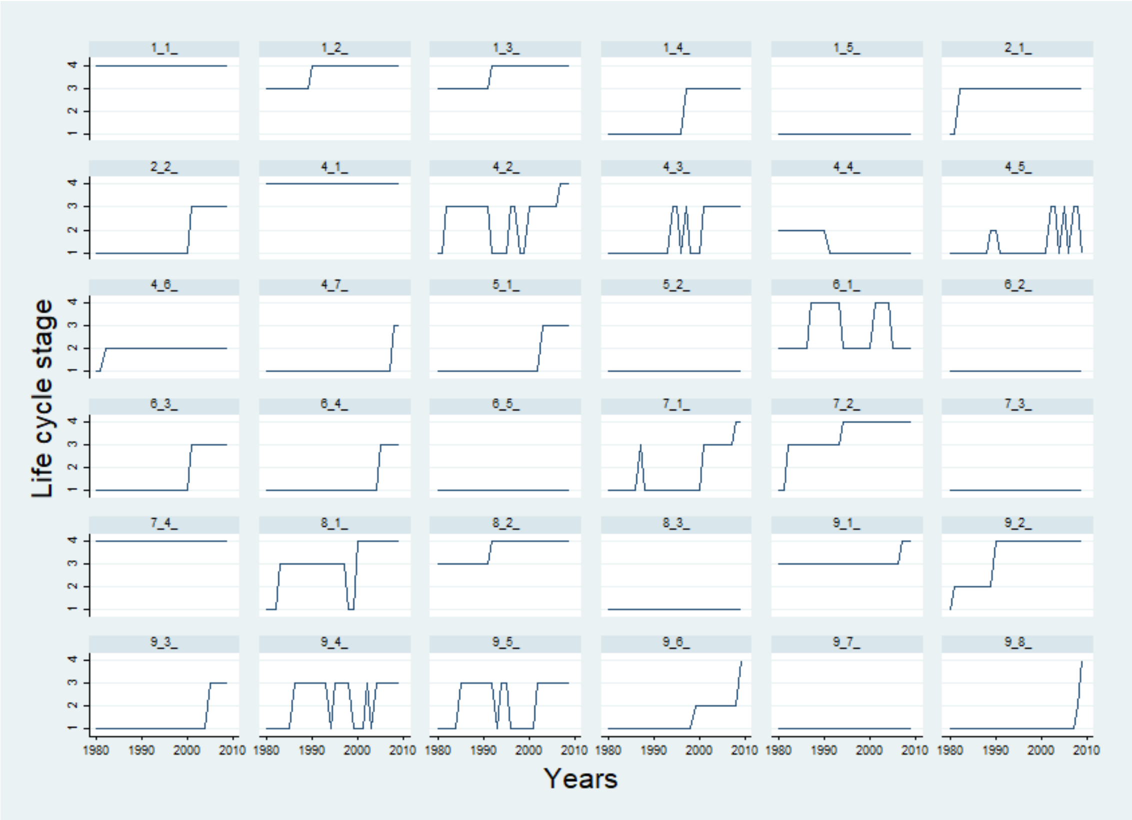


Figure A1. The life cycle of green technologies (1980-2009)

Notes: Technology names are provided in Table 2 of the paper. For the sake of space the figure reports the two-digit label of Env-Tech (OECD, 2016). Numbers in the y-axis correspond to the technology life cycle stages: 1 "Emergence", 2 "Development", 3 "Diffusion" and 4 "Maturity" (see Table 1 for a taxonomy).

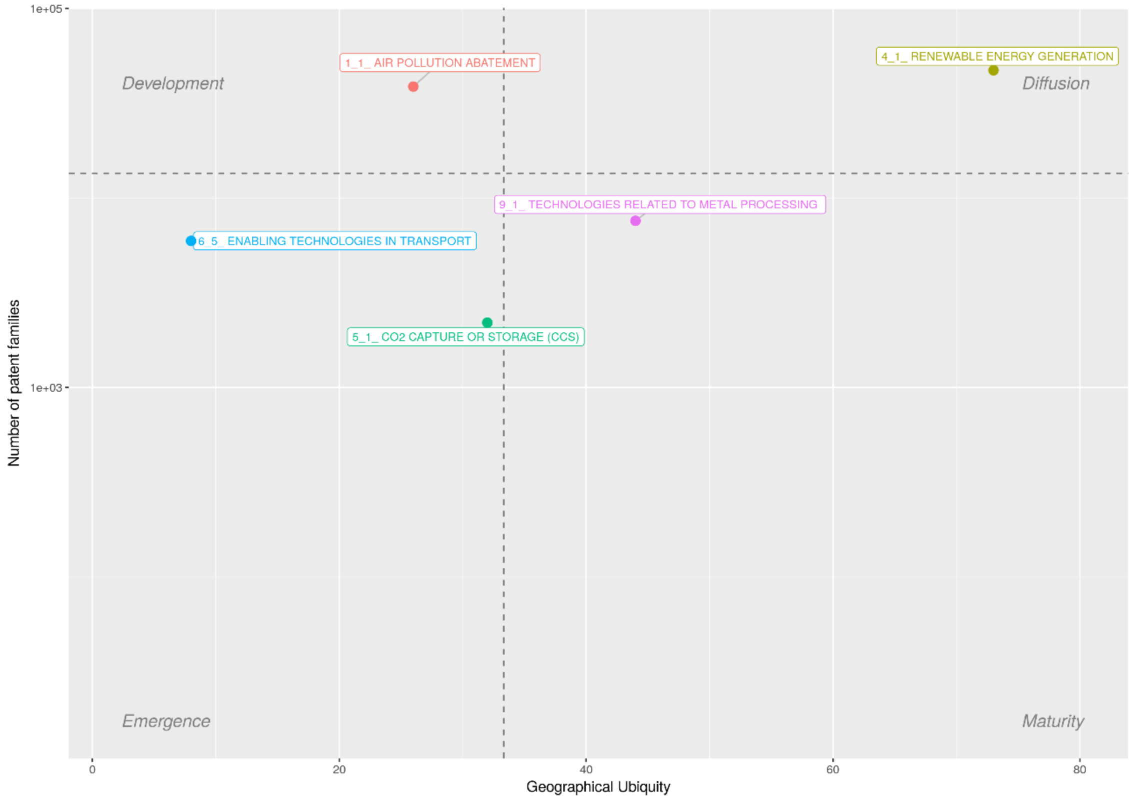


Figure A2. Selected Green Technologies by stage of life-cycle, 2000-2009

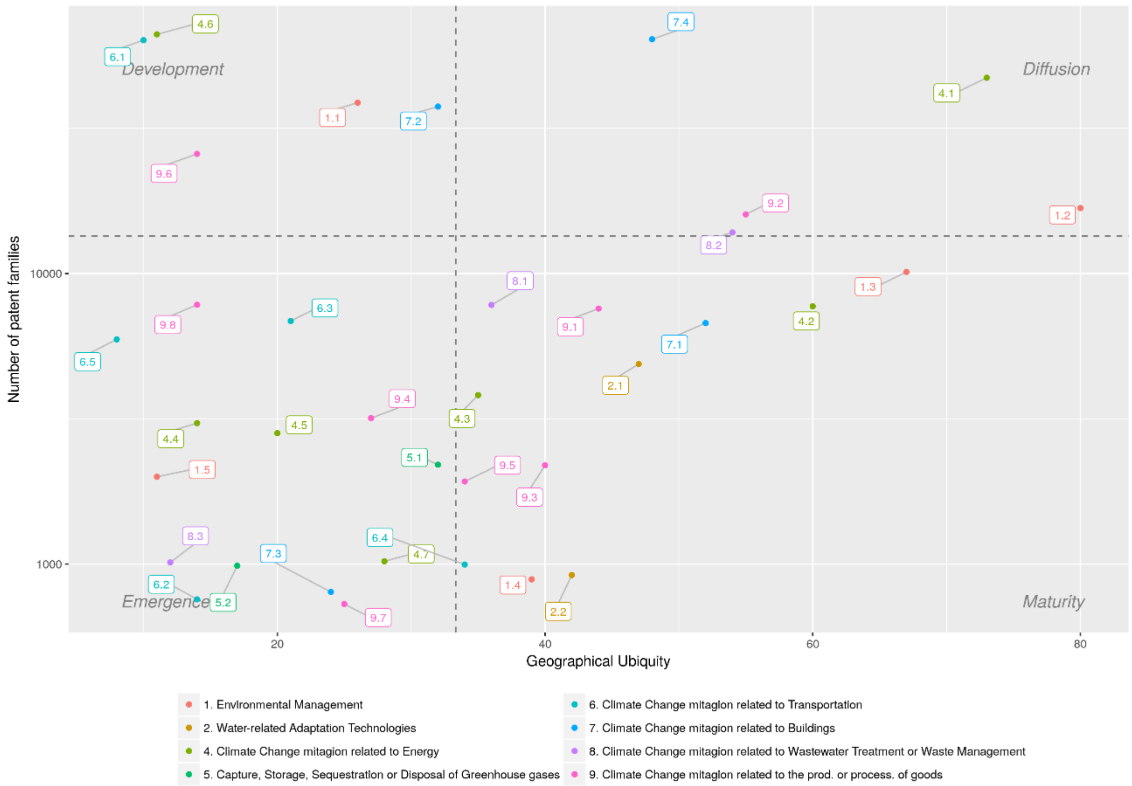


Figure A3. All green technologies by stage of life cycle, 2000-2009

Notes: Technology names are provided in Table 2 of the paper. For the sake of space, the figure reports the two-digit label of Env-Tech (OECD, 2016). Source: Own elaboration

APPENDIX C (Online publication) – Robustness checks on the technology life cycle indicators

This appendix provides robustness checks on the relationship between the knowledge base diversification and green patenting activity along the technology life cycle (TLC). In Section 3.4 and Appendix B it has been shown that the TLC indicator is built using a combination of patenting intensity and ubiquity. Following the normalization of these indicators using the standard score, (2-digit Env-Tech) green technologies were assigned to TLC stages using the mean value as a threshold to separate each stage.

We acknowledge that technologies located close to the mean value may have been characterized by an ambiguous assignment to the TLC stage. Accordingly, we check whether these technologies may have driven the main insights that arise from our empirical analysis. We identified this bunch of technologies that require a robustness check as those within a ±0.1 standard deviation from the mean value. In such a way we obtained a sample of 187 out of 1080 (36 2-digit Env-Tech for 30 years) which corresponds to 17.3% of all the green technologies considered in the present study. In order to check the assignment of these technologies to the TLC stage, we employed a k-Nearest Neighbour (k-NN) classifier. First, the k-NN has been trained using a "training" sample of 893 (1080-187) technologies that were far away from the mean value. Second, we define the "test" sample as those 187 unlabeled technologies close to the mean value. The unlabeled "test" sample is classified by assigning the most frequent TLC stage among the k=5 training samples nearest to that query point.

We find that 17 out of 187 may be assigned to another TLC stage due to their closeness to the mean value. For these technologies we employ the new assignments based on the k-NN algorithm and rerun the regressions of Table 5. The results in Table C1 show that our main findings are robust to the new assignments: regional diversification across unrelated technological domain is associated with patenting activities in technologies in an emerging phase, whereas as far as we move to maturity, unrelated variety loses importance in favour of a more related diversification. The only difference is that the coefficient of related variety in the diffusion phase is no longer significant, but this does not change the main interpretation.

Table C1. Regression results using k-NN technique to assign green technologies to life cycle stages

	GP (log)	Emergence	Development	Diffusion	Maturity
UV (log)	1.541***	1.038**	0.980	0.838	0.958*
	(0.469)	(0.497)	(0.654)	(0.912)	(0.482)
SRV (log)	0.295	-0.355	0.868***	0.134	-0.298
	(0.186)	(0.276)	(0.196)	(0.251)	(0.176)
RV (log)	0.414**	0.563*	0.470**	0.400	0.559***
	(0.168)	(0.286)	(0.216)	(0.285)	(0.112)
R&D (log)	0.0242	0.0715**	0.00668	0.0551**	-0.0226
	(0.0156)	(0.0286)	(0.0486)	(0.0233)	(0.0210)
HC (log)	0.0668	-0.118	0.289**	0.139*	0.0126
	(0.0562)	(0.0826)	(0.133)	(0.0767)	(0.0586)
R&D Neighb (log)	0.0317	0.0931**	0.158*	0.131	-0.0730
	(0.0617)	(0.0447)	(0.0817)	(0.0816)	(0.0437)
HC Neighb (log)	0.285	-0.244	0.541	-0.207	0.810***
	(0.210)	(0.332)	(0.926)	(0.247)	(0.269)
Pop Dens	0.868**	1.957***	1.079***	1.318***	0.449
	(0.370)	(0.261)	(0.260)	(0.357)	(0.515)
Inspections	0.0844	-0.0701	0.201**	-0.105	0.0558
	(0.0500)	(0.0767)	(0.0825)	(0.0655)	(0.0525)
Energy Intensity	-0.00326	0.0115	-0.0223*	-0.000123	0.00743
	(0.00627)	(0.0125)	(0.0128)	(0.00556)	(0.00458)
State FE	Y	Y	Y	Y	Y
Time Dummies	Y	Y	Y	Y	Y
Random growth	Y	Y	Y	Y	Y
Obs.	1466	1394	1365	1426	1451
R2	0.892	0.597	0.805	0.669	0.901
F	2039542.3	42825.9	22523.5	553767.3	836393.0

*Notes: The analysis covers 49 US Federal States over 1980-2009. Driscoll and Kraay's (1998) standard errors, robust to heteroskedasticity and serial and spatial correlation, in parentheses. Green technologies close to the mean value of ubiquity and patenting intensity indicators are assigned to the technology life cycle stage using k-Nearest Neighbour classification technique * p < 0.1; **p < 0.05; *** p < 0.01.*