Papers in Evolutionary Economic Geography

14.13

Relatedness in eco-technological development in European regions

Martijn van den Berge & Anet Weterings

Utrecht University
Urban & Regional research centre Utrecht

Relatedness in eco-technological development in European regions

Martijn van den Berge & Anet Weterings

PBL Netherlands Environmental Assessment Agency

The Hague, the Netherlands

P.O. Box 30314, 2500 GH The Hague, The Netherlands

Email corresponding author: Martijn.vandenBerge@pbl.nl

Abstract

Within the smart specialisation programme, the European Commission urges regional policy-makers to assess their regional innovation potential and consider investing in the areas of eco-technologies taking into account the regions' specific strengths and weaknesses. In evolutionary economic geography, several studies have shown that regional innovation is a path dependent process whereby new technologies develop out of the existing regional knowledge base. In this paper, we examine to what extent this is also the case for eco-innovation; if so, the existing technological structure of a region would be an important source of information for regional policymakers with respect to designing their eco-innovation policy agenda. Our results show that in EU-regions both the probability of developing eco-innovations and the number of patents in this field depends on the patents that have been developed in related fields in the region in prior years.

Keywords: relatedness, technology space, regional branching, eco-technologies, EU

JEL codes: C23, R11, Q55

1. Introduction

In line with what organisations such as the OECD, UNEP and World Bank have recently promoted[1], the European Commission (EC) emphasises the importance to connect sustainable and economic goals in the Europe 2020 strategy (EC 2012a). In the EU innovation policy, this ambition to stimulate sustainable growth is connected to the concept of 'smart specialisation'. The objective of smart specialisation is to boost regional innovation in order to achieve economic growth and prosperity by enabling regions to focus on their strengths[2]. To connect smart and sustainable growth, the EC explicitly encourages policymakers to assess the innovation potential of their region in areas of eco-technologies (EC 2012b).

The design of region-specific policies focused on developing eco-technologies requires policymakers to identify to what extent their region's existing knowledge base is suitable for developing eco-technologies. This is, however, not an easy task, in particular for eco-technologies. Not only are many of the eco-technologies still in an early stage of development making it highly uncertain what type of knowledge is required to develop innovations in this field, but eco-innovations also take place in a wide range of areas each building on different types of knowledge. Therefore, the ambition of the EU to develop smart and sustainable growth strategies for every EU-region asks for an analytical framework that provides insights in differences in regional capabilities to develop (different types of) eco-technologies.

Recent analyses in economic geography on 'regional branching' can contribute to the development of such a framework. The main assumption of these studies is that due to the cumulative and path dependent nature of knowledge production regions are more likely to diversify in technologies or industries that are related to their existing regional knowledge base (Boschma & Frenken 2011). New industries grow out of the existing industrial structure in a region, either out of one existing industry or through the combination of knowledge from different industries present in that region (Hildalgo et al., 2007; Neffke et al., 2011; Boschma et al., 2013a). Several recent studies showed that the same goes for technological diversification of regions (Colombelli et al. 2012; Rigby 2012; Boschma et al. 2013b; Feldman et al. 2013; Kogler et al. 2013; Rigby 2013).

Following these studies, it can be expected that European regions differ in their capabilities to develop (different kinds of) eco-technologies due to the accumulation of different types of knowledge in the past. Case study evidence already points in this direction (Cooke, 2012), yet the influence of prior technological expertise accumulated in regions on the development of eco-technologies has not been considered in a systematic way. Therefore, this paper aims to analyse to what extent this is the case for a broad range of eco-technologies being developed across Europe. Using patent statistics, we 1) measure the accumulation of technological knowledge within European regions since 1978 and 2) the extent to which this previously built-up knowledge within regions influences the development of eco-technologies within European regions.

We contribute to the few existing empirical studies on technological diversification of regions in two ways. First, we focus specifically on eco-technologies. Previous studies either studied the technological diversification of regions in general (Colombelli et al., 2012; Rigby, 2012; Boschma et

[1] www.oecd.org/greengrowth, http://org/greeneconomy, www.unep.org/greeneconomy
[2] http://ec.europa.eu/research/regions/index_en.cfm?pg=smart_specialisation

al., 2013b) or focused on one specific eco-technology (Tanner, 2011). Eco-technologies can be found in many areas of technology and, therefore, are not part of one dedicated section within the technology classification (Veefkind et al., 2012). However, a recently developed tagging system for new technological developments makes it possible to trace the different kinds of technological knowledge – as defined in the standard classification system - that are used to develop a particular eco-technology. This allows us to explore to what extent the general patterns in technological diversification also apply to these new fields of technology and, at the same time, to explore possible differences in this effect within the broad range of eco-technologies that can be identified.

Second, we not only examine whether the existing knowledge base of regions affects the probability that a region enters a new eco-technological field, but also to what extent it influences the subsequent success of the region within that field. In this latter analysis, we explore the effect of both the presence of related technological knowledge and knowledge within the specific field of eco-technology under consideration. This way we compare the relevance of technological diversification and the built up of highly specialised knowledge for further technological development within regions.

The outline of the paper is as follows. The following section describes more extensively why regional innovation processes tend to be path dependent and how the existing structure of a region may affect the likelihood of developing new technologies. Section 3 describes the data and the model and section 4 shows the results. We estimate both the probability that an eco-technology will be developed within a region and the number of patents a region may develop in the different fields of eco-technology. In section 5 we provide the conclusions and discussion.

2. Theoretical background

2.1 Path dependence in regional innovation

Innovation is often viewed as either a radical process that takes place independent from an existing stock of knowledge or as an incremental, evolutionary process where the creation of new knowledge results from deepening and re-combining the existing stock of knowledge (Sahal, 1981; Dosi, 1982; Nelson and Winter, 1982; Abernathy and Clark, 1985; Clark, 1985). In the first case, the presence of an existing knowledge stock is a poor predictor for the future innovative trajectory of firms or regions as new technological developments are radically different from previous innovations. From a locational perspective, this implies that it is unpredictable where a particular radical innovation will be developed because it does not depend on pre-existing knowledge sources within firms and territories. In other words, the windows of locational opportunity are completely open (Storper and Walker, 1989).

However, there is strong evidence to believe this radical nature of innovation is rather the exception than the rule (Simmie, 2012). Like all actors, firms are characterised by bounded rationality, not having the capabilities to select the best and most promising or profitable technologies. High risks and switching costs prevent them to build completely new technologies from scratch (Breschi et al., 2003; Boschma et al., 2013; Makri et al., 2013), rather they focus on technology domains which present similarity in problem solving and knowledge bases (Nelson and Winter, 1982; Dosi, 1997). As a result, new technologies generally do not emerge in virgin markets but merely are the result of an incremental process where new capabilities are related to pre-existing capabilities (Witt, 2003; Neffke et al., 2011). In essence, it is claimed that technological change is not random but path dependent (Perez, 2010).

This more incremental view on innovation can be translated to a locational perspective as well. Numerous studies have pointed out that knowledge spillovers – i.e. through formal and informal networks, spin-offs or labor market mobility – are at least partly non-tradable and spatially bounded, thereby taking place within limited geographical place[3]. As firms and other organisations depend on these external knowledge sources in their local environment they are often what is called 'locally embedded' (Frenken and Boschma, 2007; Buenstorf and Klepper, 2009).

Connecting the idea of path dependency and local embeddedness, several studies have put forward countries as an important unit of analysis in studying the cumulative and path dependent nature of economic change (Hildalgo et al., 2007; Hausmann and Hildalgo, 2010). Moreover, economic geographers have claimed that the regional scale might be even more important for this process of diversification (Boschma et al., 2013a), as 'localized capabilities' are regional intangible assets with a high degree of tacitness that are difficult to replicate in other places, even within countries (Neffke, 2009). Following this view, it can be argued that 'the windows of locational opportunity' are not completely 'open'. Hence, we expect that the creation of new eco-technologies in regions is mostly related to a set of existing capabilities.

2.2 Regional knowledge structures and innovation capacity

The economic geographic literature traditionally views the level of sectoral specialisation as one of the most important factors that affects the potential of regions to develop new technologies. There is a widespread discussion whether regions benefit mainly from a set of highly specialised capabilities or from a variety of capabilities, also referred to as respectively 'Marshallian externalities' (localisation economies) and 'Jacobs' externalities' (urbanisation economies). In this line of reasoning differences in regional innovation and growth are related to qualitative differences in an economy's composition at the regional level (Frenken et al., 2007). Localisation economies follow from a strong specialisation in specific sectors or technologies creating cost-reducing externalities due to the better matching of skilled labour and input-output transactions, and more effective learning by means of knowledge spillovers. The high degree of cognitive proximity that exists between actors in those regions facilitates knowledge exchange between economic agents (Cohen and Levinthal, 1990; Nooteboom, 2000; Boschma, 2005). Urbanisation economies, on the other hand, develop within regions with more variety in technological knowledge which enables the cross-fertilization of ideas among sectors, and in that way, generates more inventions.

However, more recent studies argued that this distinction between localisation and urbanisation economies is too limited. Not an economic structure characterised by a high level of specialisation or

[3] Mechanisms explaining why knowledge spillovers are spatially bounded are for example the necessity of having face-to-face contact, the importance of informal social networks and its localness for knowledge exchange and the limited geographical scope of labor market mobility (see Boschma (2005) for a critical overview on the role of spatial proximity for innovation).

diversity, but 'related variety' is the most advantageous for innovation and economic growth (e.g. Frenken et al., 2007; Neffke et al., 2011; Rigby, 2012; Boschma et al., 2013a and b). The high degree of cognitive proximity which is assumed to trigger Marshallian externalities could actually have an adverse effect on the innovative capacity of the region on the long run caused by a highly specialised, unilateral knowledge stock. To put it plainly, on the long run there is nothing new to learn from one another. And Jacobs' externalities are assumed to arise from a somewhat spontaneous cross-fertilization between a variety of cognitive distant capabilities; a process that is too much dependent on serendipity and, consequently, limited in explaining economic and technological change.

Related variety refers to the variety of industries present in a region that are cognitively related (Frenken et al., 2007). When the degree of related variety is higher in a region, more learning opportunities are available at the local level, and consequently, more knowledge spillovers across industries are likely to occur. The local presence of a wide range of technologically related industries provides local learning opportunities and growth potentials for existing industries as well as local sources of growth for new industries. In this latter respect, related variety may spur technological diversification and true economic renewal in regions through increasing the probability that new recombinations between related technologies will occur.

2.3 Case study evidence for technological diversification in eco-technologies

Some case study evidence points in the direction that also eco-technological development is the result of recombining existing regional capabilities. Cooke (2008; 2012) describes for example how the clean tech industries in northern and southern California evolved from the geographical convergence in agro-food, ICT and biotechnology in that region, Jutland's wind and solar thermal clusters in Denmark from the specialisation in agricultural equipment and marine engineering, and Wales' photovoltaic and biofuel technology development from its mining equipment and agro-food clusters. Using a more systematic analyses of patent data on fuel cell technology development within Europe, Tanner (2011) that specific technologically related knowledge fields are significantly co-located with the generation of fuel cell development. The higher the number of related technological fields present in a given region, the more likely a region is to branch into fuel cell technology.

The study by Fornahl et al. (2012) takes a more critical stand. They examined the anecdotal assumption that the emergence of the offshore wind energy industry in northern Germany was strongly related to pre-existing capabilities in the shipbuilding industry. Their study showed that only a few offshore wind energy firms have their roots in shipbuilding. Primarily it were onshore wind energy firms or firms of other industries, partly from other regions, that have diversified into offshore wind energy. The yard crisis and shipbuilding decline did provide human capital and competences in steel construction, electronics in autonomous systems, maritime logistics and the handling of heavy weight components which could be used by offshore wind energy firms. However, according to Fornahl et al. (2012) similar conditions could be found in other regions with different industrial structures. Human capital as such cannot be seen as sufficient evidence for new path creation out of established paths. The development of the industry was more likely to result from the basic locational condition of access to seaports and related infrastructure, positive market development and the reaction of national and governments on the yard crisis and shipbuilding decline by increased support investment schemes, R&D spending and infrastructure investments.

2.4 Hypotheses on regional branching of eco-technologies

Following the literature described above, we expect that regions in which the knowledge base is characterised by a high level of knowledge that is related to a certain eco-technology are (1) more likely to branch into developing that eco-technology and, subsequently, (2) that those regions have a higher innovation output in those technological fields. In other words, the emergence of eco-technologies in regions and the following success of regions in developing those innovations depends on the existing set of technological capabilities within regions.

3. Data and method

3.1 Data

To test whether the presence of (a combination) of technologies related to eco-technologies within a region affects both the emergence and success in developing eco-technologies in those regions, we use patent data from the OECD REGPAT databases (July 2013) including patent applications filed to the European Patent Office (EPO) from 1977 to 2009. Patent statistics have been widely used in quantitative studies to measure levels of technological competence for units of analysis ranging from the individual to the regional level. These statistics encompass a wealth of information to investigate knowledge creation and diffusion processes (Pavitt, 1985; Grilliches, 1990; Frenken et al., 2007). In general, three main advantages can be identified. First, patent data contain highly detailed information on content (title, words, abstracts and technologies involved), inventors (names), organisation(s) involved (institutional affiliation) and geographical location (addresses). Second, systematic data collection goes back a long time and, third, the current 'stock' of patents is extensive and continues to expand.

Despite these advantages, using patent data to measure knowledge production has also been widely criticized. We should bear in mind that their use is subject to limitations. More specifically, three major drawbacks can be identified (Grilliches, 1990). First, research does not necessarily lead to patents. Rejection is one of the main reasons. Other reasons include the time and cost constraints of researchers with regard to the submission of reports for patenting, and the non-disclosure strategies of firms who value secrecy more highly than they value property rights. Moreover, patents are codified knowledge, whereas a high proportion of knowledge produced in firms, universities, and research institutes are tacit (Patel and Pavitt, 1997). Second, patents do not necessarily contribute to our knowledge. Most patents are rarely cited, if at all, suggesting that they add little value to the knowledge system and the commercial value of patents also varies widely. Third, patenting rates differ systematically across scientific disciplines and technological fields, respectively. Consequently, differences in technological specialisation can therefore render inter-regional comparisons misleading.

Nevertheless, for the purpose of this study, patent applications are considered to be the most appropriate measure given their relative homogenous, detailed, and consistent recording of knowledge production. Eco-technological development is an economy-wide transformation (Frankhauser et al., 2012), and consequently, data on firm establishments or industries are not useful for analysing the development of these technologies. Moreover, recently a new tagging system has been developed for eco-technologies (we will further elaborate on this system below) making patents the most suitable data for defining eco-innovation in a systematic manner.

To analyse regional differences in technological development, we use the REGPAT database in which patent data has been linked to regions utilising the addresses of the applicants and inventors. Similar as in other studies on regional patent data, we use the link based on the inventor's address, since this is considered to be closest to the place of invention. REGPAT covers more than 5,500 regions across OECD countries, EU27 countries, Brazil, China, India, the Russian Federation and South Africa. The analyses in this paper are limited to the European regions. Due to low numbers of patent activity in eastern European countries, we focus on 16 countries (EU15 countries, minus Greece and plus Norway and Switzerland). In total, we included 202 NUTS2 regions in the analysis (regions in all 16 countries excluding the island regions of France, Spain, Italy, Portugal and Finland[4]).

Eco-technologies are relatively new technological areas that are not (yet) included in the existing international patent classification (IPC) system that EPO uses to reflect the scope of the approved claims listed in a patent document. To still be able to identify eco-patents we used the general tagging system of new technological developments (Y02 codes) of the Cooperative Patent Classification (CPC). Appendix A provides an overview of the selected Y02 classification[5].

As we are interested in the relatedness of eco-technologies to other technological fields it is important to choose an appropriate level of aggregation containing enough homogeneity within one eco-technology. At the 5-digit level, not all Y02 classes are homogenous enough for the purpose of our analysis. For instance, the class Y02E1 "energy generation through renewable energy sources" still contains a substantial variation in technological knowledge since it consists of geothermal energy (Y02E1/1), hydro energy (Y02E1/2), energy from the sea (Y02E1/3), solar thermal energy (Y02E1/4), photovoltaic (PV) energy (Y02E1/5), thermal-PV hybrids (Y02E1/6), and wind energy (Y02E1/7). Therefore, we decided to define eco-technologies at the 6-digit level. We excluded overlapping codes such as photovoltaic technology applied in buildings (Y02B classes) and photovoltaic energy technology in general (Y02E classes). In this case we excluded the first one as it is a more detailed classification of the latter[6]. This results in a distinction between 35 eco-technologies.

---

[4] FR83, FR91, FR92, FR93, FR94, ES63, ES64, ES70, ES83, ITG1, ITG2, PT20, FI20.

[5] http://www.uspto.gov/web/patents/classification/cpc/html/cpc-Y.html. Y codes have been originally created by the EPO (European Classification System, ECLA) as an extension of the original classification system, to extend classification capabilities to new technology areas of special interest, and are usually added in addition to codes in the A through H series. The Y classification was created specifically to cover "General Tagging of New Technological Developments". Subjects covered by ECLA Y codes originally included nanotechnology topics. In 2010, a new set of Y codes were added for clean energy technologies. ECLA has been replaced (including Y code classification) by the Cooperative Patent Classification (CPC) in 1 January 2013. This is a bilateral system which has been jointly developed by the EPO and the USPTO.

[6] We measured overlap by the number of co-occurrences of Y02 codes within patents.

3.2 Composing the technology space

Similar as Boschma et al. (2013a) for all patents in the US, we composed a *technology space* to measure to what extent eco-technologies are related to the IPC technology classes. Building on the method developed by Hidalgo et al. (2007), who composed a so-called *product space*, that is, a network representation of the relatedness between products traded in the global economy. In a similar vein, we we constructed a network-based representation of the relatedness between all the technologies that can be found in the patent portfolio of the EPO database[7]. In this one mode *n*n* (IPC*IPC) network each node *i* represents a specific technological class and edges connect nodes to each other when related. How stronger technologies (nodes) are related to one another, how closer they are positioned in the network. In general, technologies with a high degree of relatedness are positioned more central in the network and more isolated technologies more peripheral.

To compose this network, we computed the relatedness between each IPC class *i* with every other IPC class *j* at the 4-digit level by taking the normalized co-occurrence (size-effects) within all patents (comparable to Breschi et al., 2003). We adopted a probabilistic similarity measure - the association strength - as proposed by Van Eck & Waltman (2009). The probabilistic approach resulted in a measure of the co-occurrence of technology *i* and *j* that equals:

$$(1)\quad \phi_{ij} = \frac{o_{ij}}{e_{ij}}$$

Where ϕ_{ij} stands for the relatedness between technology *i* and *j* ; o_{ij} for the observed co-occurrence between *i* and *j* ; and e_{ij} for the expected co-occurence. This measure has a straightforward probabilistic interpretation: $o_{ij} / e_{ij} > 1$, when *i* and *j* co-occur more frequently than would be expected by chance and $o_{ij} / e_{ij} < 1$ when *i* and *j* co-occur less frequently than would be expected by chance. We computed the expected co-occurrences as follows:

$$(2)\quad e_{ij} = \left(\frac{s_i}{p} * \frac{s_j}{p} \right) * p$$

Where s stands for the number of times technology *i* and *j* are occurring in the REGPAT database during the period in consideration and p stands for the total number of documents (patents) in the same period.

In one formula:

$$(3)\quad \phi_{ij} = \left(\frac{o_{ij}}{s_i s_j} \right) * p$$

[7] We also used the patent portfolio of the selection of EU countries and found a correlation of the relatedness measures >0.9. We decided to include all patents to construct the technology space (not only our selection of EU countries) as we have no arguments to add location in this step of the analyses, as it is a place-neutral relatedness network.

Confidence intervals were constructed using a binomial test to determine which co-occurences are significantly lower or higher than the expected co-occurences with a p-value of 5% (see for more details on this method Neffke and Henning, 2009). Finally, we kept only the combinations with a relatedness index >1, i.e. observations which co-occur significantly more than expected. A visual impression of the technology space based on all patents in the period 1978-2009 is provided in Appendix 2.

Since eco-technologies are not captured in the IPC, we had to separately construct the technology space for each eco-technological class Y (at the 6-digit level) with all other IPC-classes (at the 4-digit level). This resulted in a two mode m*n (Y*IPC) matrix, all the previous methods of measuring relatedness being the same. In some cases, this lead to problems concerning the assumption of statistical independence. For example, in the case of wind energy (Y02E1/1) we find a high relatedness with "wind motors" (F03D), that is probably not caused by a high degree of relatedness, but due to an overlap in the classifications, that is, almost all wind energy patents also contain the wind motor classification (F03D) but not all wind motor patents contain the code for wind energy. We dealt with this problem by counting the number of related technologies within regions after excluding all wind energy patents. Only those "wind motor patents" were counted which have not been classified as being a "wind energy patent".

Another deviation from the total IPC*IPC matrix is that in the Y*IPC matrix we only took into account IPC subclasses (4-digit level) that have a share of >1% in the total of co-occurrences with a given eco-technology Y (6-digit level). A practical reason for this is that limited number of occurrences of a particular Y and IPC code, let's say *i* and *j*, sometimes leads to a probabilistic outcome of a co-occurrence between *i* and *j* that is significantly more than expected, even if the absolute number of co-occurrences is very low. Sometimes even a co-occurrence of 1 turned out to be significantly more than expected. In Appendix 3 we visualize in the total technology space which IPC codes are related to three examples of eco-technologies (average of period 1978-2009).

Note that we used all available patents in the REGPAT database to compute the relatedness matrix, that is, also patents from countries outside our sample of European countries. In particular in the early period of observation, the number of patents in certain technology classes is very low. To avoid too large fluctuations in the relatedness of technologies, we, therefore, used a 9 year moving average to determine which technologies are related to eco-technologies starting in 1978, the second year of available data in the REGPAT database, thereby covering the years 1978-1986 for the first period. A consequence of using a moving average for measuring technological relatedness is that our model estimations are limited to patents in eco-technologies in the period 1982-2005. In 1982 we used information from the relatedness matrix for the period 1978-1986 - covering 4 years prior and 4 years after the year of observation - and 2005 is the last possible year that can be included in the analyses as for that year we used information from the relatedness matrix of 2002-2009. Furthermore, we only included eco-technologies (Y02 codes at the 6-digit level) with more than 100 patents in our selection of European regions during the time period of the analysis (1982-2005). Following our selection of countries, regions and the time period, in total 1,227,621 patents of the REGPAT database were included in our analyses of which 31,257 in eco-technologies.

3.3 Measuring the variables

Two dependent variables

We composed two dependent variables to measure the effect of the existing regional knowledge base of regions on both the probability that a region starts to develop eco-technologies and how successful it is in this field. The first dependent variable is called ‘entry’ and has a score of 1 for the first year in which a region has a share of patents in one of the eco-technologies that is higher than the European average and 0 for all other years[8]. This is called a revealed comparative advantage (RCA) of the region. The second dependent variable measures the success of a region in developing an eco-technology. This variable measures the number of patents in a certain field of eco-technology that each region has developed for each year of observation.

The patents are ascribed at the regional level using a non-fractional count. Although the OECD REGPAT database provides fractional counts in case there are several inventors with different regional residences who were involved in developing the invention, we share the view of Tanner (2011) that knowledge is a non-divisible asset. Thus, if multiple inventors from different regions are involved in developing a patent we assign the patent to each region involved.

Independent variables

Relatedness density in regions

We measured the presence of technologies related to the different eco-technologies within regions using the information from the ‘technology space’ as defined in section 3.2 We excluded all eco-patents from the database before aggregating the frequency of patents within the related IPC classes by region, thereby calculating the number of patents in related technologies on the base of all patents that are not assigned as an eco-patent (see also Tanner, 2011). This is done to circumvent the problem of measuring the production of eco-knowledge instead of knowledge related to eco-knowledge (we illustrated this problem before using the example of “wind energy” and the relatedness with “wind motors”). We took the cumulative counts of patents in related technology classes for all previous years, and following Zucker et al. (2007) discounted this number by 20% annually to reflect deprivation of knowledge. We repeated this for every technology *i* and the total number of patents. These cumulative counts were used to calculate the revealed comparative advantage (RCA) for each technology class in a given region as follows:

$$RCA_{i,r,t} = \frac{p_{r,t}(i) / \sum_i p_{r,t}(i)}{\sum_r p_{r,t}(i) / \sum_r \sum_i patents_{r,t}(i)} > 1$$

Next, we computed the relatedness density measure as developed by Hildalgo et al. (2007) and recently adopted in patent statistics by Boschma et al. (2013). This measure indicates how close a

[8] Note that having a revealed comparative advantage (RCA) for the first time often equals entering with just one patent for the first time due to low numbers of patents in eco-technologies, in particular, in the earlier period.

new technology is to the existing technological portfolio of a given region. The density around a given technology *i* (in this case eco-technology *y*) in region *r* is computed from the technological relatedness of technology *i* to the technologies in which the region *r* has a RCA in time *t*, divided by the sum of technological relatedness of technology *i* to all the other technologies in the EPO patent portfolio:

$$RD_{i,r,t} = \frac{\sum_{j \in r, i \neq j} \phi_{ij}}{\sum_{i \neq j} \phi_{ij}} \times 100$$

With this measure, we test whether eco-technologies are more likely to be developed in regions that have a knowledge base with a larger variety of technologies that are related to eco-technologies. This increases the likelihood of recombining the different technologies that are necessary for developing eco-technologies. To avoid too large year-by-year fluctuations in the relatedness density of a region, we took the average over the five year period prior to the year of observation.

Average relatedness

The advantage of the relatedness density measure is that we count the degree of variety around a given eco-technology in regions. The disadvantage is that we do not take into account the importance of an IPC class in terms of its degree of relatedness with a particular eco-technology. To test whether our findings differ when we do take those differences into account, we also calculated the so-called average relatedness around a given eco-technology per region. This measure is introduced by Feldman et al. (2013) and calculated as follows:

$$AR_{i,r,t} = \frac{\sum_j \phi_{ij}^t * S_{j,r,t}}{P_{r,t}}$$

Where ϕ_{ij}^t represents the relatedness between an eco-technology *i* and IPC classes *j* in time *t*. $S_{j,r,t}$ stands for the number of patents in a region *r* in time *t* in related technologies *j*. $P_{r,t}$ is the total number of patents in region *r* in time *t*. Compared to the RD measure, an important disadvantage of this measure is that a high presence of one related IPC class can already result in a high value of AR even when all other related IPC classes are not or poorly at present in a given region.

Control variables

Regions which are generally more active in patenting than other regions are also more likely to develop a patent in eco-technology in a certain year. Therefore, we control for the total number of patents being developed by regions in every year (cumulative counts for all previous years with a discount factor of 20% annually). In addition, we included euros spent on research and development activities (R&D expenditures) to control for the general level of knowledge development in each region. We also included population density to control for externalities related to urbanization economies and the number of inhabitants (population) to control for differences in regional size. In the models in which we estimated the number of eco-technologies being developed in regions every year, we also included a control variable to account for the effect of the number of eco-patents applied in the foregoing years. This variable is composed in a similar way as the count of patents in related technologies and total patents, i.e., cumulative counts for all previous years, and discounting by 20% annually to reflect depreciation of knowledge. In this way we are also able to test whether the relatedness density around a particular eco-technology is still important after regions increasingly specialize in the eco-technology itself.

At the technology level we adopt three control variables, as suggested by Boschma et al. (2013). The first variable is the size of the technology measured by the number of inventors involved. One can expect that higher inventive activity in a particular technology increases the probability that regions develop patents in this technology. The second variable is the year-by-year growth rate of patents to account for the expansion of technological opportunities. We expect that the growth in patenting will enhance the probability of regions to start developing patents in that technology as well and in the number of patents being developed. The third variable is the concentration of a technology. The more the patenting activity in a technology is regionally concentrated, the less likely it will that other regions develop patents[9]. Finally, we included regional, technology and time fixed effects by including dummy variables for each of these three groups. The first two groups are included to control for any time-independent regional and technological characteristics that may correlate with the two main explanatory variables. Time fixed effects are used to control for yearly differences in patenting activity that are not region or technology specific.

3.4 Models

We used two types of models to examine how the regional knowledge base affects regional patenting in eco-technologies: a discrete time duration model to estimate the event that a region develops an above average share of patents in an eco-technology in a certain year (RCA > 1) and a count model to estimate the number of patents that a region develops once this technology emerged in this region.

Discrete time duration models are used to model time-to-event data when the event may take place at any point in time but no information is available on the exact moment of the event (Jenkins, 2005). The REGPAT database reports the number of patents per region on a yearly basis and, consequently, it is only possible to observe changes in the development of eco-technologies from one year to another while the actual event could have taken place at any moment during that year. The dependent variable, the time spell from 1982 to the first time that a region develops an above average share of patents in any of the eco-technologies defined is left censored, as for most eco-technologies, several regions already developed a patent before 1982. For data with such a structure, duration analysis is the most appropriate methodology (Guo, 1993).

The methodology that was adopted to model the event of developing a RCA in eco-technology is the complementary log-logistic (cloglog) function which is the most commonly-used discrete time representation of a continuous time proportional hazards model (Jenkins, 2005). This model

[9] See for formulas Boschma et al. (2013)

essentially tells you how likely it is that a region will develop an above average share of patents in an eco-technology in a certain year, given that it has not yet done so until that year. By specifying dummy variables to represent each year, the baseline hazard rate has been modeled as a step function that describes the evolution of the baseline hazard between censored intervals. For further technical details regarding discrete time duration models and, more specific, the complementary log-logistic function, see Jenkins (2005).

We included all regions in the entry model, but a region is dropped from the panel database once it enters the field of eco-technologies by developing a RCA larger than 1 in that field. Four regions not yet developed any patents in eco-technologies in 2005[10]. These regions were dropped from the analyses once we included regional fixed effects in the model.

In the second model, we estimated the number of patents in eco-technologies that has been developed in a region in a certain year. This variable consists of integers which never have negative values and our data is clearly characterized by overdispersion. Since many regions also did not develop any eco-patents in particular years of the observed time period, a zero inflated negative binominal model is the most appropriate[11]. This regression equation has two components: a regime selection equation (inflate) and a count data component. The regime selection equation estimates how likely it is that there will be any patent developed at all in the region, that is, the probability of regions to have a value of zero. Next, the count data component estimates the number of patents developed, assuming that a non-zero regime is selected. In these models, we only included those regions that have at least developed once a patent in one of the selected eco-technologies as this allows us to include fixed effects on the regional level.

The control variables at the regional level have the same score for all 35 eco-technologies and the control variables at the technology level have the same score for all 202 regions - only the relatedness density measure is region and technology specific. To avoid a bias from estimating the effects of those aggregated explanatory variables, models were estimated with cluster-robust standard errors on the regional level (Steenbergen and Jones, 2002)[12].

Table 1 presents the summary statistics. The variance inflation factors show that multicollinearity did not pose a problem.

---

[10] ES23, ES43, PT15, PT18

[11] We tested whether our data is characterized by zero inflation using fit statistics (Voung, BIC, AIC) to compare the fit of a negative binomial model with that of a zero inflated negative binomial model. All fit statistics showed that the latter regression model has a better fit.

[12] Since the central question of this paper is how relatedness density affects the development of eco-technologies in regions after controlling for other regional characteristics, it was considered to be more important to estimate the model with cluster robust standard errors on the regional level. Tests with clustering on technology level showed that such a control does not change our results.

Table 1 *Descriptive statistics and variance inflation factors (VIF)*

Entry

Variable	Level**	N	Min	Max	Mean	S.D	VIF
Entry RCA in eco-patents	REG	121,992.00	0.00	1	0.03	0.16	
Relatedness Density (% t-1) †	REG	121,992.00	0.00	100	22.17	18.22	1.17
Average Relatedness (% t-1) †	REG	121,992.00	0.00	97.77	2.5	2.71	1.00
R&D expenditure (€ mil. t-1) *	REG	121,992.00	0.00	14,670.77	354.34	543.05	1.93
Total patents (t-1) *†	REG	121,992.00	0.00	33,306.63	965.69	1,647.45	1.78
Population (1.000 t-1) *	REG	121,992.00	106.96	11,319.97	1,555.08	1,244.53	1.52
Population Density	REG	121,992.00	0.00	9.13	0.32	0.76	1.07
Geographical concentration	TECH	121,992.00	0.00	27.96	0.59	1.52	1.02
Number of inventors*	TECH	121,992.00	0.00	2,478.00	137.04	215.51	1.15
Technological growth rate	TECH	121,992.00	-1.00	0.99	0.08	0.23	1.11

Count

Variable	Level**	N	Min	Max	Mean	S.D	VIF
Number of eco-patents	REG	84,696.00	0.00	57.00	0.33	1.34	
# eco-patents previous yrs. (t-1) †	REG	84,696.00	0.00	158.32	1.11	3.75	1.77
Relatedness Density (% t-1) †	REG	84,696.00	0.00	100.00	32.05	20.92	1.07
Average Relatedness (% t-1) †	REG	84,696.00	0.00	36.45	2.99	2.47	1.01
R&D expenditure (€ mil. t-1) *	REG	84,696.00	0.00	14,670.77	856.40	1,382.94	1.77
Total patents (disc.fact t-1) *†	REG	84,696.00	0.00	33,306.63	2,661.17	4,037.78	1.66
Population (1.000 t-1) *	REG	84,696.00	106.96	11,319.97	2,126.73	1,620.86	1.67
Population Density	REG	84,696.00	0.00	9.13	0.48	1.01	1.04
Geographical concentration	TECH	84,696.00	0.00	27.96	0.63	1.02	1.02
Number of inventors*	TECH	84,696.00	0.00	2,478.00	245.23	341.67	1.11
Technological growth rate	TECH	84,696.00	-1.00	0.99	0.10	0.20	1.05

**= log-transformed in model estimations. Variance inflation factor (VIF) based on log-transformed variables.*
*** =REG are variables measured on the regional level, TECH on the technological level*
†= Based on cumulative patent counts discounted 20% annually to reflect deprivation of knowledge.

Finally, we run several models as a robustness check. First, we tested whether our results change when we replace relatedness density by the alternative measure average relatedness. Second, we excluded the top and bottom 10 regions in patent activity from the sample to test whether our results are sensitive for extreme cases. Finally, we run models for each eco-technology separately to test whether the effect of relatedness density in eco-innovations differs per eco-technology. There is a wide range of eco-technologies and possibly some innovations are more radical making relatedness density less relevant in those cases. In these models, we included fixed effects on the country instead of the regional level, because many regions never developed any patents in one of those fields during the time of observation. We did include country level fixed effects because large differences exist between European countries in the extent to which the development of these three eco-technologies are stimulated which may cause country level differences in the number of patents in those fields.

4. Empirical results

Table 2 shows the results of the complementary log-log models that we used to estimate the probability that a region developed a share of patents in a certain eco-technological field in a certain year that is higher than the European average for that technology at that time, given that it has not yet developed a RCA larger than one in the years before. Model 1 shows that the relatedness density of a region in a prior period indeed increases the entry probability of that region in an eco-technology, as shown by the positive and statistically significant effect of this variable. In other words, regions that already have developed a RCA in fields that are related to the specific eco-technology are more likely to also develop a RCA in that eco-technology.

The positive and statistically significant effect of relatedness density does not change when we add control variables at the regional and technological level in model 2. Two of the four control variables at the regional level have the expected sign: regions are more likely to enter in a certain field of eco-technology when the general patenting activity within the region is higher (total patents) and when more investments in R&D take place within the region. The effect of population density and the number of inhabitants (population) is not statistically significant indicating that the size and level of urbanisation of regions has no specific effect on the probability that regions start developing patents in eco-technologies. The three control variables on the technology level show that an overall higher number of inventors active in the technological field under consideration and a higher growth rate of patents in that field both increase the probability that a region develops an above average share of patents in that technological class. The effect of the geographical concentration of patenting activity in a technological field has a negative effect on the probability of entry, suggesting that geographical concentration of technological development decreases the probability that a technology will be developed in a region. However, this effect is not statistically significant[13].

Table 2 *Results of complementary log-log model of probability to have a RCA in an eco-technology in a certain year (robust standard errors in parentheses – clustered on NUTS-2 level)*

	1	2	3 <=1995
Relatedness Density (RD)	0.015***	0.016***	0.019***
	(0.001)	(0.001)	(0.002)
Ln(R&D expenditure t-1)		0.109*	-0.113
		(0.063)	(0.109)
Ln(Total patents t-1)		0.508***	0.376*
		(0.109)	(0.204)
Ln(Population t-1)		-0.115	-0.921
		(0.659)	(1.304)
Population Density t-1		0.024	-0.797
		(0.314)	(0.523)
Geographical concentration t-1		-0.002	-0.011
		(0.022)	(0.036)

[13] To control for overestimation of the effect of the variables on the technology level we also estimated a model with robust standard errors at this level. This caused no changes in the significance levels of the all variables.

Ln(Number of inventors t-1)		0.358***	0.370***
		(0.042)	(0.067)
Technological growth rate t-1		0.443***	0.221
		(0.099)	(0.139)
Constant	-4.830***	-9.458*	-1.284
	(0.158)	(4.918)	(9.670)
Regional fixed effects	YES	YES	YES
Technology fixed effects	YES	YES	YES
Time fixed effects	YES	YES	YES
Observations	118,632	118,632	73,340
Nonzero observations	3,057	3,057	1,571
Wald ^2	3,817***	3,994***	2,536***
Log Likelihood	-12294	-12206	-6328

*** p<0.01, ** p<0.05, * p<0.1

To provide further insights in the relevance of relatedness density of the regional knowledge base for developing patents in eco-technologies, we calculated by which percentage the base hazard rate will increase when moving from the 50th to the 75th percentile which is an increase of about 15% in relatedness density. If the level of relatedness density for a given eco-technological field in a given region would change to such an extent, the probability that this region will develop an above average share of patents in this field increases by 24 percent. This is a substantial effect, although it is somewhat lower than what Boschma et al. (2013b) found for technologies in general in the US[14].

In model 3, we excluded all years after 1995. After this year patenting activity in most eco-technologies grew rapidly and, therefore, the period up to 1995 can be considered as the early stage of development. For this period, the effect of relatedness density is also positive and statistically significant and considerably higher: moving from the 50th to the 75th percentile in relatedness density in a given technology and region increases the probability of developing an above average share of patents in that field by 34 percent. The effect of R&D expenditures and technological growth both turn statistically insignificant in model 3.

The second question we raise in this paper is whether relatedness density of the regional knowledge base affects the success of European regions in patenting in the different eco-technology fields once patenting activity in that field took off in the region. Table 3 shows the results of the zero inflated negative binomial models that we used to estimate the number of patents in eco-technologies being developed in European regions in a year between 1982 and 2005. In the model, we included all regions where at least once patenting in eco-technology has taken place during the period of observation[15].

In the regime selection equation (inflate), we included relatedness density and regional control variables. We find a negative and statistically significant effect of relatedness density (see model 2 in Table 3). This confirms our expectation about the effect of this indicator because a negative effect here indicates that a higher relatedness density decreases the probability that a region belongs to

[14] When Boschma et al. (2013b) included fixed effects in the models, the rate of entry increased by approximately 30 percent for a 10 percent increase in the level of density in city-technology pairs.

[15] While the exclusion of ES23, ES43, PT15, PT18 in the entry model was because these region developed no patents in any of the eco-technologies, we now judge per technology if they have developed any patent and if not, we exclude these regions for these technologies (for other technologies they are kept in the analyses).

the zero-regime, that is, that the region is unlikely to develop any patents in the eco-technology under consideration. We also find a negative and statistically significant effect of the total patent activity within the region and the region's population; the more populated and the higher the general patenting rate in regions, the higher the number of patents in eco-technologies developed there. R&D expenditures and population density, on the contrary, both have a positive effect in this part of the model, however, this effect is in both cases not statistically significant.

The count data component of model 1 in Table 3 shows that, similar as in the entry model, relatedness density of European regions has a positive and statistically significant effect on the number of patents in eco-technologies being developed in a certain year. In other words, regions that have developed a RCA in more technological fields that are related to the different eco-technologies are not only more likely to develop an above average patenting activity in that field but those regions are also more likely to continue to develop more patents in that eco-technology.

When we add the control variables on the regional and technological level in model 2 in Table 3, the coefficient of relatedness density becomes smaller, but the effect remains positive and statistically significant. Model 2 shows that the number of patents in an eco-technology being developed in a region also depends on how many patents that region has developed in previous years, both in general (total patents) and specific in that eco-technological field. This is shown by the positive and statistically significant effect of those two regional level variables. Regions that were already more active in the field and in patenting in general are more likely to stay active in this field later on. In line with the results of the entry model, the number of inventors and the growth rate of patenting in the different technological fields has a positive effect on the number of eco-patents developed in European regions, while technological concentration has a negative but not statistically significant effect.

Table 3 *Results of the zero-inflated negative binomial model estimating the number of patents developed in each field of eco-technology and in each region, per year (robust standard errors in parentheses – clustered on NUTS-2 level)*

	1	2	3 >=1995
Relatedness Density (RD)	0.017***	0.007***	0.010***
	(0.002)	(0.001)	(0.002)
Ln(Patents in eco-tech t-1)		0.768***	0.734***
		(0.022)	(0.033)
Ln(R&D expenditure t-1)		0.020	-0.335**
		(0.029)	(0.166)
Ln(Total patents t-1)		0.323***	0.594***
		(0.097)	(0.186)
Ln(Population t-1)		-0.458	-1.173
		(0.493)	(1.433)
Population Density t-1		0.304*	-0.532
		(0.181)	(0.509)
Geographical concentration t-1		-0.024	-0.024
		(0.020)	(0.037)
Ln(Number of inventors t-1)		0.282***	0.251***

		(0.030)	(0.043)
Technological growth rate t-1		0.781***	0.576***
		(0.071)	(0.106)
Constant	-2.795***	-2.774	2.462
	(0.187)	(3.719)	(10.508)
Regional fixed effects	YES	YES	YES
Technology fixed effects	YES	YES	YES
Time fixed effects	YES	YES	YES
Inflate			
Relatedness Density (RD)	-0.007	-0.018***	-0.025
	(0.008)	(0.005)	(0.027)
Ln(R&D expenditure t-1)	0.093	0.023	0.021
	(0.093)	(0.060)	(0.493)
Ln(Total patents t-1)	-0.625***	-0.539***	-0.244
	(0.118)	(0.094)	(0.249)
Ln(Population t-1)	-0.611*	-0.632**	-0.763
	(0.337)	(0.252)	(0.872)
Population Density t-1	0.097	0.080	-0.824
	(0.129)	(0.073)	(3.590)
Constant	6.886***	7.349***	6.145
	(1.684)	(1.184)	(4.574)
LnAlpha	0.277***	-0.394***	-0.485***
	(0.089)	(0.116)	(0.183)
Observations	84,696	84,696	45,877
Wald ^2	12,757***	17,177***	5,830***
Log Likelihood	-46790	-44580	-15824

*** p<0.01, ** p<0.05, * p<0.1

The results in table 3 raise the question whether the patenting activity of a region in a certain field of technology depends more on the development of technological knowledge in the related fields of technology or in the eco-technological field itself. To test this we calculated for both variables how an increase comparable to a move from the 50^{th} to the 75^{th} percentile affects the number of patents being developed in a given eco-technological field and given region. Such an increase in the number of patents already developed in the field of eco-technology within the region during previous years raises the number of patents being developed later on by 54%, while a similar increase in relatedness density leads to 11% more patents in that eco-technological field. The strong effect of the number of patents in eco-technology developed in the past suggests a learning effect: once regions are active in that field of eco-technology they are likely to continue developing more patents in the same field. When we limit our analysis to the period 1995-2005, both variables still have a positive and statistically significant effect (see model 3 in Table 3). During that time period, the effect of the number of patents in the own eco-technological field is somewhat smaller (44% by an increase from the 50^{th} to the 75^{th} percentile), but still substantially larger than the effect of relatedness density by a similar increase (16%).

Robustness check

As a robustness check (Table 4) we first replaced relatedness density by average density, the alternative measure of relatedness used by Feldman et al. (2013). Similar as what we found for relatedness density, average density has a positive and statistically significant in both the entry and the count model (model 1 and 2 respectively in Table 4). However, the effect of average density in the regime selection equation of the count model is positive instead of negative. In other words, regions with a higher average density are less likely to develop any patents in the eco-technology under consideration, but if they do start developing these patents they develop more patents in this field.

In models 3 to 6 shown in Table 4, we tested whether our results where not driven by extreme cases. We do so by excluding the upper top 10 regions in total patent activity in model 3 (entry) and 4 (count) and the bottom 10 regions in model 5 (entry) and 6 (count). Excluding these cases does not influence the direction and significance of the coefficients in the entry (model 3 and 5) and the count equation of the zero inflated negative binominal model (model 4 and 6). Only in the regime selection of model 4, the negative coefficient of relatedness density becomes statistically insignificant.

Finally, we estimated separate entry and count models for each of the 35 eco-technologies to see whether relatedness density has a positive effect on patenting in each of those fields, as these technologies differ quite substantially in both patenting activity and required technology. The results are shown in appendix 4. The table shows that indeed for the majority of the 35 eco-technologies relatedness density has a positive effect on both the probability that a region develops an above average patenting rate in that field (entry) and the number of patents that a region develops in that field (count). Nevertheless, in some cases, relatedness density has no statistically significant effect on the probability of entry (e.g., energy from sea, solar thermal energy, PV energy, fuel from waste and energy storage). The same goes for the count models. Case study research such as the study by Fornahl et al. (2013) would be very helpful in developing a better understanding of why this is the case.

Table 4 *Robustness check (robust standard errors in parentheses – clustered on NUTS-2 level)*

	1 ENTRY	2 COUNT	3 ENTRY <top10	4 COUNT <top10	5 ENTRY >bottom10	6 COUNT >bottom10
Average Density (AD t-1)	0.103***	0.104***	-	-	-	-
	(0.010)	(0.008)				
Relatedness Density (RD t-1)	-	-	0.017***	0.010***	0.016***	0.007***
			(0.001)	(0.002)	(0.001)	(0.001)
Ln(patents in eco-tech t-1)	-	0.783***	-	0.740***	-	0.769***
		(0.021)		(0.026)		(0.022)
Ln(R&D expenditure t-1)	0.127*	0.009	0.076	0.014	0.105*	0.013
	(0.066)	(0.028)	(0.058)	(0.036)	(0.064)	(0.028)
Ln(total patents t-1)	0.588***	0.283***	0.453***	0.400***	0.529***	0.302***
	(0.098)	(0.098)	(0.108)	(0.114)	(0.115)	(0.100)
Ln(population t-1)	-0.340	-0.483	-0.023	-0.540	-0.130	-0.407
	(0.609)	(0.480)	(0.660)	(0.535)	(0.671)	(0.495)
Population Density t-1	-0.103	0.115	0.006	0.272	0.014	0.301*

	(0.363)	(0.218)	(0.322)	(0.183)	(0.316)	(0.179)
Geographical concentration t-1	0.009	-0.021	0.004	-0.005	-0.002	-0.024
	(0.021)	(0.020)	(0.023)	(0.023)	(0.022)	(0.020)
Ln(Number of inventors t-1)	0.344***	0.264***	0.388***	0.381***	0.360***	0.281***
	(0.041)	(0.028)	(0.045)	(0.035)	(0.043)	(0.030)
Technological growth rate t-1	0.472***	0.786***	0.460***	0.776***	0.423***	0.784***
	(0.100)	(0.071)	(0.110)	(0.102)	(0.101)	(0.072)
Constant	-8.013*	-2.165	-9.764**	-3.394	-9.466*	-2.905
	(4.595)	(3.635)	(4.936)	(4.100)	(5.027)	(3.734)
Regional fixed effects	YES	YES	YES	YES	YES	YES
Technology fixed effects	YES	YES	YES	YES	YES	YES
Time fixed effects	YES	YES	YES	YES	YES	YES
inflate						
Average Density (AD t-1)		0.079***		-		-
		(0.023)				
Relatedness Density (RD t-1)		-		-0.017		-0.020***
				(0.029)		(0.005)
Ln(R&D expenditure t-1)		-0.033		0.054		0.028
		(0.050)		(0.093)		(0.062)
Ln(total patents t-1)		-0.772***		-0.450		-0.626***
		(0.097)		(0.278)		(0.103)
Ln(population t-1)		-0.457		-1.218		-0.531**
		(0.299)		(0.824)		(0.265)
Population Density t-1		0.237***		-0.186		0.059
		(0.051)		(0.698)		(0.078)
Constant		6.802***		9.300*		7.317***
		(1.470)		(5.029)		(1.266)
LnAlpha		-0.353***		0.163**		-0.397***
		(0.111)		(0.079)		(0.116)
Observations	118,632	84,696	113,237	70,128	105,900	82,968
Nonzero observations	3,057		2,982		9,068	
Wald ^2	3,931***	3,427***	3,568***	17,745***	8,421***	16,880***
Log Likelihood	-12237	-44598	-10834	-30700	-11801	-44148

*** p<0.01, ** p<0.05, * p<0.1

5. Conclusions and discussion

This paper investigated the effect of the presence of (a variety of) eco-technology related technologies in European regions on the development of eco-technologies in those regions between 1982 and 2005. In general, the results show that regions with a knowledge base in technological fields that are related to different kinds of eco-technologies are 1) more likely to entry in eco-technologies and, subsequently, 2) more successful in developing patents in those eco-technologies. In other words, this study confirms the assumption of the literature on related variety and regional branching that technological development within regions depends on the existing set of capabilities in those regions, not only in general (Boschma et al., 2013a), but also in particular for eco-technologies.

These results show that information about the relatedness between technologies can be an important input for the design of regional policies aimed at stimulating smart and sustainable growth within regions, as it provides a better understanding of the potential within regions for innovation and technological diversification. Similar as what has been shown for new industries (Neffke et al., 2011), new technologies also grow out of the existing technological structure in a region, either out of one existing technology or through the combination of knowledge from different technologies present in that region. Consequently, which eco-technologies are most likely to successfully develop within a region largely depends on the type of technologies being developed in that region until now, as regional economic structures tend to develop through a process of related diversification (Boschma & Frenken, 2011). Following this idea, it is useless to pursue 'one-size-fits-all' policies or try to develop new economic structures from scratch (Boschma, 2013).

Although the analyses in this paper provide a clear indication of the relevance of the existing set of regional capabilities for developing eco-technologies, there are several options to further extend our insights in this process. First, it is important to further test whether our findings depend on the way we measure technological relatedness. Leten et al. (2007) and Rigby (2012), for instance, have used alternative indicators to measure the relatedness between patent classes, such as patent citations.

Second, it would be good to add an institutional perspective. We only included technological variables in our models to explain eco-technological development within regions and did not pay any attention to institutional factors such as differences in environmental policies and regulations (Porter and Van der Linde, 1996), policies about the internalization of environmental costs (Acemogle et al., 2012; Naoily and Smeets, 2012), subsidies, differences in attitudes towards climate change or in entrepreneurial attitudes, innovation policies and the spatial scale they are at play (country or region). Such factors can results in substantial differences between countries and regions in the probability and number of patents in eco-technologies being developed.

A third recommendation is to add information on interregional connections. The analyses in this paper are limited to characteristics of the region itself, while links with other regions may also affect the probability and subsequent success of regions in developing new technologies. While the presence of related knowledge is highly important for the absorptive capacity of regions, besides cognitive proximity also geographical or social proximity may contribute to the diffusion of new technologies across regions (for empirical evidence see for instance Feldman et al., 2013). Also from a policy perspective, insights in the importance of such interregional links are interesting. Possibly, such links could offer regions with limited related variety in certain technologies the opportunity to circumvent their lack of knowledge through establishing (long distance) interregional networks.

Fourth, as it is very hard to distinguish between incremental and radical innovations, we were not able to test whether the development of some eco-technologies may be the result of crossovers between unrelated technologies. This is an interesting question for further research as it is a highly relevant question from a policy perspective. Policies which focus on stimulating breakthrough innovation probably need a different design than policies focusing on exploiting related diversification. Policy concepts like the smart specialization strategy may be blind to the potential of unrelated diversifications, while it could be argued that regions might need to make a jump into more unrelated activities now and then (Boschma, 2013). It might be fruitful to connect such regional policies to other policy concepts with more focus on stimulating breakthrough innovations like a strategic niche management (SNM) approach which aims to create a niche environment outside the context of the existing technological regime (Simmie, 2012).

Finally, future research should further explore which mechanisms are underlying the process of technological diversification and regional branching. In this paper, we examined these processes on the regional level, while they actually follow from micro-level behavior such as firm diversification, the establishment of spin-offs, labor mobility or knowledge exchange between firms, universities or other research institutes (see Breschi & Lissoni, 2004; Tanner, 2014). Those microlevel processes should be further studied to get a good understanding of how technological diversification within regions actually takes place and how policies can stimulate those processes.

References

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

Acemoglu, D., Aghion, P., Bursztyn, L. and David, H. (2012) The environment and directed technical change. American Economic Review 102(1): 131-166.

Boschma, R.A. (2005) Role of proximity in interaction and performance. Conceptual and empirical challenges, Regional Studies, 39(1): 41-45.

Boschma, R.A. and Frenken, K. (2011) The emerging empirics of evolutionary economic geography. Journal of Economic Geography, 11(2): 295-307.

Boschma R. and Frenken K. (2011) Technological relatedness and regional branching, In: H. Bathelt, M.P. Feldman and D.F. Kogler (eds.) Beyond Territory. Dynamic Geographies of Knowledge Creation, Diffusion and Innovation, pp. 64-81. London and New York: Routledge.

Boschma, R.A. (2013) Constructing regional advantage and smart specialisation: comparison of two European policy concepts. Papers in evolutionary geography. Urban & Regional Research Centre Utrecht, Utrecht University.

Boschma, R., Balland, P.A. and Kogler, D. (2013a) Relatedness and Technological Change in Cities: The rise and fall of technological knowledge in U.S. metropolitan areas from 1981 to 2010. Industrial and Corporate Change, forthcoming.

Boschma, R., Minondo, A. and Navarro, M. (2013b) The emergence of new industries at the regional level in Spain. A proximity approach based on product-relatedness, Economic Geography, 89(1): 29-51.

Buenstorf, G. and Klepper, S. (2009) Heritage and agglomeration: the Akron tyre cluster revisited, The Economic Journal, 119: 705-733.

Breschi, S. and Lissoni, F. (2001) Knowledge spillovers and local innovation systems: a critical survey, Industrial and Corporate Change 10: 975–1005.

Breschi, S., Lissoni, F., and Malerba, F. (2003) Knowledge-relatedness in firm technological diversification. Research Policy, 32(2003): 69–87.

Breschi, S. and Lissoni, F. (2004) Knowledge networks from patent data: methodological issues and research targets. KITeS Working Paper #150, KITeS, Centre for Knowledge, Internationalization and Technology Studies, Universita' Bocconi, Milano, Italy.

Castaldi, C., Frenken, K., and Los, B. (2013) Related Variety, Unrelated Variety and Technological Breakthroughs: An analysis of U.S. state-level patenting. Papers in evolutionary economic geography, Utrecht University.

Clark, K.B. (1985) The interaction of design hierarchies and market concepts in technological evolution, Research Policy 14(5): 235–251.

Cohen, W.M. and Levinthal, D. A. (1990) Absorptive capacity: a new perspective on learning and innovation, Administrative Science Quarterly 35: 128-152.

Colombelli A., Krafft, J., Quatraro, F. (2012) The emergence of technology-based sectors at the regional level: a proximity-based analyses of nanotechnology. Papers in evolutionary economic geography, Utrecht University.

Cooke, P. (2008) Regional Innovation Systems, Clean Technology & Jacobian Cluster-Platform Policies. Regional Science Policy & Practice 1(1): 23-45.

Cooke, P. (2012) Transversality and transition: green innovation and new regional path creation. European Planning Studies 20(5): 817-834.

Dosi G. (1982) Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change, Research Policy 11(3): 147–162.

EC (2012a) Connecting smart and sustainable growth through smart specialisation. S3 Smart specialisation platform.

EC (2012b) Guide to research and innovation strategies for smart specialisation (RIS 3). European Commission, RIS3 Smart specialisation Platform. Brussels, May 2012.

Eck, N.A. Van and Waltman, L. (2009) How to normalize cooccurrence data? An analysis of some well-known similarity measures. Journal of the American Society for Information Science and Technology 60(8): 1635-1651.

Feldman, M.P., Kogler, D.F., and Rigby, D. (2013) rKnolwedge: the spatial diffusion of rDNA Methods. Working paper, UCLA Department of Geography.

Fornahl, D., R. Hassink, C. Klaerding, I. Mossig and Schröder, H. (2012) From the old path of shipbuilding onto the new path of offshore wind energy? The case of northern Germany. European Planning Studies 20(5): 835-855.

Frankhauser, S., A. Bowen, R. Calel, A. Dechezleprêtre, D. Grover. J. Rydge and Sato, M. (2012) Who will win the green race? In search of environmental competitiveness and innovation, working paper no. 11 London: Centre for Climate Change Economics and Policy, working paper no.94 London: Grantham Research Institute on Climate Change and the Environment.

Griliches, Z. (1990) Patent statistics as economic indicators. A survey, Journal of Economic Literature 28: 1661-1707.

Guo, G. (2003) Event-history Analysis for Left-truncated Data, Sociological Methodology, 23: 217-243.

Hägerstrand T. (1953), Innovation Diffusion as a Spatial Process. Lund: University of Lund.

Heimeriks, G. and Boschma, R. (2013) The path- and place-dependent nature of scientific knowledge production in biotech 1986-2008, Journal of Economic Geography, forthcoming.

Hausmann, R. and Hidalgo, C.A. (2010) Country diversification, product ubiquity, and economic divergence, CID Working Paper No. 201, Center for International Development, Harvard University.

Hidalgo, C.A., Klinger, B., Barabasi, A.-L. and Hausmann, R. (2007) The product space conditions the development of nations, Science 317: 482–487.

Jenkins S.P. (2005) Survival analysis, unpublished manuscript. Institute for social and economic research, University of Essex, Colchester, UK, downloadable from: http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/pdfs/ec968lnotesv6.pdf

Jørgensen, U. and Karnøe, P. (1995) The Danish wind turbine story: technical solutions to political visions, in A. Rip, T. Misa & J. Schot (eds.) Managing Technology in Society: the Approach of Constructive Technology Management, London, Pinter.

Kemp R., Schot, J. and Hoogma, R. (1998) Regime shifts to sustainability through processes of niche formation: the approach of strategic niche management. Technology analysis & strategic mangament 10(2): 175-196.

Kogler, D.F., Rigby, D. and Tucker, I. (2013) Mapping Knowledge Space and Technological Relatedness in US Cities. European Planning Studies 21(9): 1374-1391.

Leten B., Belderbos, R. and Van Looy, B. (2007) Technological diversification, coherence, and performance of firms, Journal of Product Innovation and Management 24: 567-579.

Makri, M., M.A. Hitt, and P.J. Lane (2010) Complementary technologies, knowledge relatedness, and invention outcomes in high technology mergers and acquisitions. Strategic Management Journal 31(6): 602-628.

McCann, P. (2013) Place-based development policies and smart specialisation. In: Thissen, M., F. van Oort, D. Diadato and A. Ruijs (2013), Regional competitiveness and smart specializiation in Europe. Edward Elgar Publishing.

Neffke, F. (2009) Productive Places. The Influence of Technological Change and Relatedness on Agglomeration Externalities. PhD thesis, Utrecht University, Utrecht.

Neffke, F., Henning, M. (2009) Skill-relatedness and firm diversification. Working paper #0906 of Papers on Economics and Evolution. Jena: Max Planck Institute of Economics, Evolutionary Economics Group.

Neffke, F., Henning, M. and Boschma, R (2011) How do regions diversify over time? Industry relatedness and the development of new growth paths in regions, Economic Geography 87: 237-265.

Noailly, J. and Smeets, S. (2012) Directing Technical Change for Fossil-Fuel to Renewable Energy innovation: an Empirical Application Using Firm-level Patent Data, The Hague: CPB Netherlands Bureau for Economic Policy Analysis.

Nooteboom, B. (2000) Learning and innovation in organizations and economies. Oxford University Press, Oxford.

Patel, P., and Pavitt, K. (1997) The technological competencies of the world's largest firms: Complex and path-dependent, but not much variety. Research Policy 26: 141-156.

Pavitt, K. (1985), 'Patent statistics as indicators of innovative activities: Possibilities and problems', Scientometrics 7: 77-99.

Perez, C. (2010) Technological revolutions and technological paradigms. Cambridge Journal of Economics 34(1): 185-202.

Porter, M.E., and Van der Linde, C. (1995) Towards a new conception of the environment competitiveness relationship. Journal of Economic Perspectives 94: 97-118.

Rigby D. (2012) The Geography of Knowledge Relatedness and Technological Diversification in U.S. Cities. Papers in evolutionary economic geography, Utrecht University.

Rigby D. (2013) Technological relatedness and knowledge space: entry and exit of U.S. Cities from Patent Classes. Working paper. University of California Los Angeles.

Schot, J. and Geels, F.W. (2008) Strategic niche management and sustainable innovation journeys: theory findings, research agenda, and policy. Technological analysis & strategic management 20(5): 537-455.

Simmie (2012) Path dependence and new technological path creation in the Danish wind power industry. European planning studies 20(5): 753-772.

Smoch (2008) Concept of a technology classification for country comparisons. Final report of the World intellectual property organisation (WIPO). Fraunhofer institute for systems and innovation research. Karlsruhe, Germany.

Steenbergen, M.R. and Jones, B.S. (2002) Modeling Multilevel Data Structures, American Journal of Political Science 46(1): 218-237.

Storper, M. and Walker, R. (1989) The capitalist imperative: Territory, technology, and industrial growth. New York: Wiley-Blackwell.

Tanner, A.N. (2011) The place of new industries: the case of fuel cell technology and its technological relatedness to regional knowledge bases. Papers in evolutionary economic geography, Utrecht University.

Tanner, A.N. (2014) Regional Branching Reconsidered: Emergence of the Fuel Cell Industry in European Regions. Economic Geography, doi: 10.1111/ecge.12055

Sahal, D. (1981) Patterns of Technological Innovation. Addison Wesley, Reading, MA.

Veefkind, V., Hurtado-Albir, J., Angelucci, S., Krachalalios, K., Thumm, N. (2012) A new EPO classification scheme for climate change mitigation technologies. World Patent Information, 34: 106-111.

Witt, U. (2003), The evolving economy. Cheltenham: Edward Elgar.

Zucker L.G., Darby M.R., Furner J., Liu R.C., Ma H. (2007) Minerva unbound: Knowledge stocks, knowledge flows and new knowledge production. Research Policy, 36: 850-863.

Appendix 1: Y02E-CLASSIFICATION

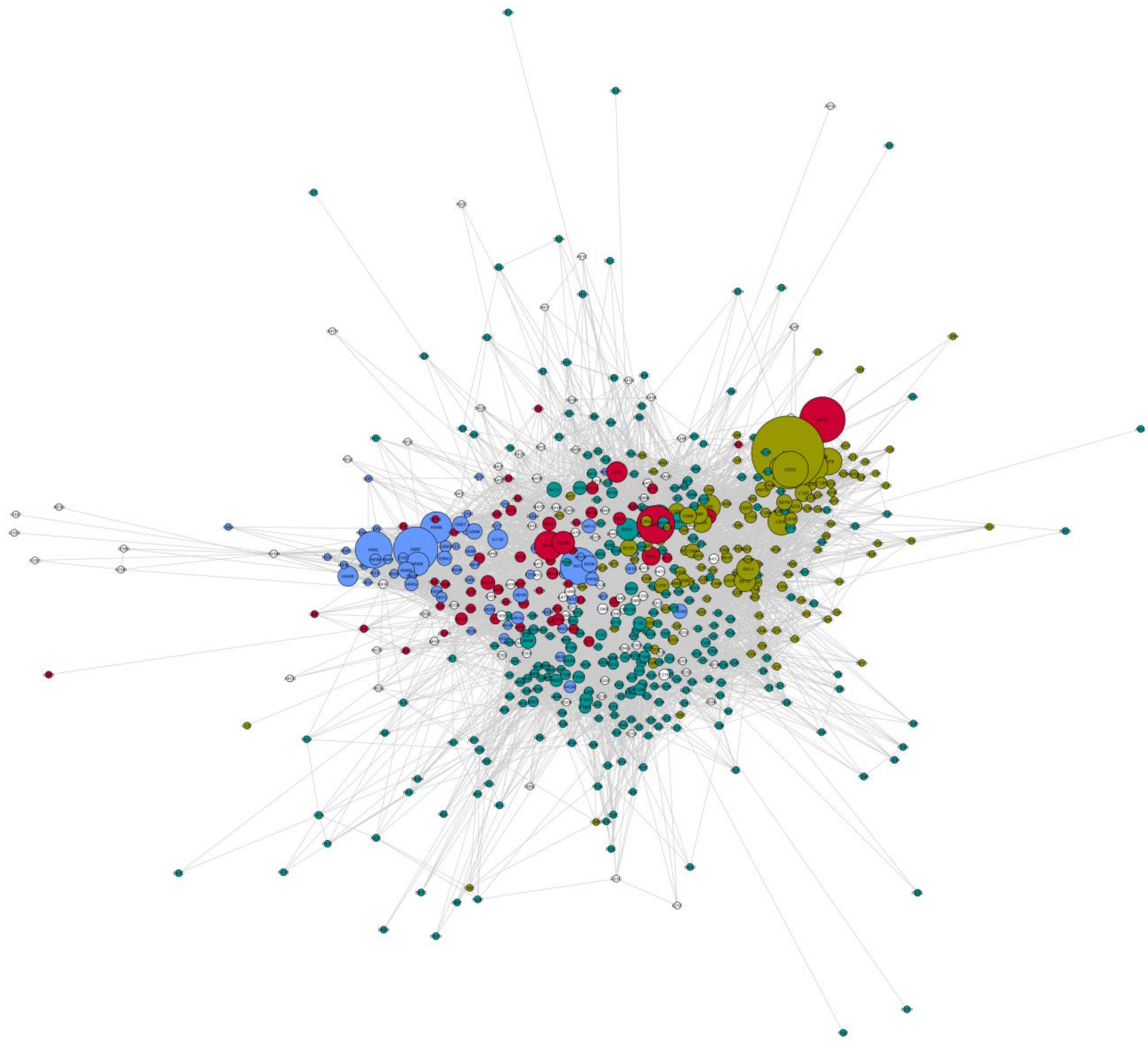
Cooperative patent classification (CPC), system for tagging new technological development for mitigation or adaptation against climate change. See for detailed information: http://www.epo.org/news-issues/issues/classification/classification.html

Table Selection of CPC patent classification system for tagging new technological development for mitigation or adaptation against climate change (Y02)

	CPC code	Description
	Y02B	*CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TO BUILDINGS*
1	Y02B 20/1	Energy saving technologies for incandescent lamps
2	Y02B 20/2	High pressure [UHP] or high intensity discharge lamps [HID]
3	Y02B 20/3	Semiconductor lamps
4	Y02B 30/1	Energy efficient heating, using boilers and heat pumps
5	Y02B 30/5	Systems profiting of external/internal conditions (e.g. heat recovery)
6	Y02B 30/6	Other technologies for heating or cooling
7	Y02B 30/7	Efficient control or regulation technologies
8	Y02B 60/1	Energy efficient computing
9	Y02B 60/3	Techniques for reducing energy-consumption in wire-line communication networks
10	Y02B 70/1	Technologies improving the efficiency by using switched-mode power supplies [SMPS]
11	Y02B 90/2	Smart grids in building environment
	Y02C	*CAPTURE, STORAGE, SEQUESTRATION OR DISPOSAL OF GREENHOUSE GASES [GHG]*
12	Y02C 10/0	CO2 Capture by biological and chemical separation and adsorption
13	Y02C 20/1	Capture of nitrous oxide (N2O)
	Y02E	*REDUCTION OF GREENHOUSE GASES [GHG] EMISSION, RELATED TO ENERGY GENERATION, TRANSMISSION OR DISTRIBUTION*
14	Y02E 10/2	Hydro energy
15	Y02E 10/3	Energy from sea
16	Y02E 10/4	Solar thermal energy
17	Y02E 10/5	Photovoltaic [PV] energy
18	Y02E 10/7	Wind energy
19	Y02E 20/1	Combined combustion
20	Y02E 20/3	Technologies for a more efficient combustion or heat usage
21	Y02E 40/6	Superconducting electric elements
22	Y02E 50/1	Biofuels
23	Y02E 50/3	Fuel from waste
24	Y02E 60/1	Energy storage
25	Y02E 60/3	Hydrogen technology
26	Y02E 60/5	Fuel cells
	Y02T	*CLIMATE CHANGE MITIGATION TECHNOLOGIES RELATED TOTRANSPORTATION*

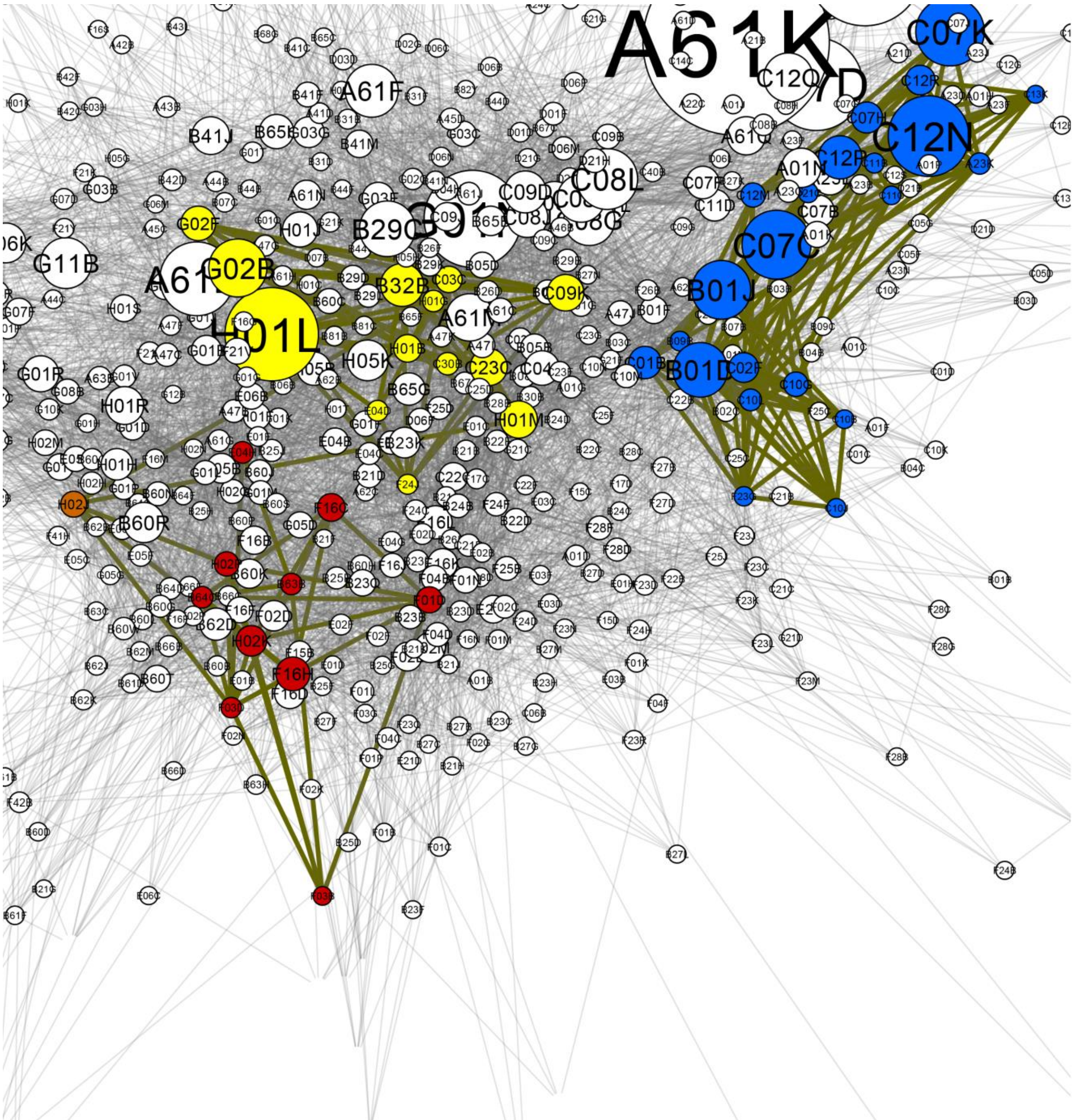
27	Y02T 10/1	Internal combustion engine [ICE] based vehicles
28	Y02T 10/4	Engine management systems
29	Y02T 10/6	Other road transportation technologies with climate change mitigation effect
30	Y02T 10/7	Energy storage for electromobility
31	Y02T 10/8	Reduce green house gasses emissions in road transportation
32	Y02T 50/1	Drag reduction
33	Y02T 50/4	Weight reduction
34	Y02T 50/6	Efficient propulsion technologies
35	Y02T 90/3	Application of fuel cell technology to transportation

Appendix 2: Technology space 1978-2009 (IPC 4-digit technologies)



Legend	
	Electrical engineering
	Instruments
	Chemistry
	Mechanical engineering
	Other fields

Appendix 3: Three examples of related IPC4 codes to eco-technologies visualized in total technology space 1978-2009



WIND ENERGY	Y02E 10/7
PV ENERGY	Y02E10/5
BIOFUELS	Y02E50/1

Appendix 4: Coefficients and standard errors of Relatedness Density measure for sample models

	ENTRY (FE**)		COUNT (FE**)					
	Relatedness Density		Relatedness density				Ln(patents in eco-tech t-1)	
			Count		*inflate*		*count*	
TECH*	coef.	S.E.	coef.	S.E.	coef.	S.E.	coef.	S.E.
1	0.021***	-0.006	0.008	-0.005	-0.038***	-0.013	0.659***	-0.192
2	0.017**	-0.009	-0.003	-0.004	-0.027**	-0.011	1.308***	-0.145
3	0.013**	-0.006	-0.004	-0.005	-0.028***	-0.008	0.967***	-0.145
4	0.017***	-0.006	0.008*	-0.004	-0.077	-0.097	0.757***	-0.076
5	0.012*	-0.006	-0.002	-0.009	-0.031*	-0.017	1.212***	-0.315
6	0.005	-0.008	-0.008	-0.01	-0.023	-0.018	1.336***	-0.174
7	0.011*	-0.006	0.009	-0.009	-0.019	-0.019	0.653***	-0.185
8	0.028***	-0.007	0.013***	-0.005	-0.017	-0.012	1.142***	-0.172
9	0.015***	-0.005	0.003	-0.004	-0.016**	-0.007	0.391***	-0.141
10	0.024***	-0.004	0.004	-0.004	-0.026***	-0.009	0.897***	-0.086
11	0.024***	-0.005	0.001	-0.004	-0.037***	-0.007	0.746***	-0.103
12	0.022***	-0.007	0.007	-0.005	-0.031***	-0.009	1.102***	-0.083
13	0.017**	-0.007	-0.003	-0.008	-0.025	-0.035	0.840***	-0.206
14	0.018***	-0.006	0.002	-0.006	-0.036	-0.052	1.218***	-0.135
15	0.001	-0.01	-0.027***	-0.01	-0.038*	-0.022	1.723***	-0.246
16	0	-0.007	0.001	-0.003	-0.02	-0.017	0.942***	-0.057
17	0.005	-0.008	-0.005*	-0.003	-0.029**	-0.012	0.945***	-0.046
18	0.020***	-0.006	0.005	-0.003	-0.089***	-0.034	1.316***	-0.058
19	0.023***	-0.007	0.008**	-0.004	-0.043***	-0.011	0.732***	-0.079
20	0.020***	-0.006	-0.002	-0.005	-0.038**	-0.016	0.909***	-0.125
21	0.014	-0.01	0	-0.004	-0.049***	-0.016	1.412***	-0.253
22	0.026***	-0.006	0.017***	-0.006	-0.007	-0.012	0.558***	-0.112
23	-0.003	-0.007	0.001	-0.004	-0.016	-0.01	0.940***	-0.102
24	0.005	-0.009	-0.001	-0.002	-0.045**	-0.019	0.900***	-0.035
25	0.018**	-0.007	-0.006***	-0.002	-0.032***	-0.009	0.783***	-0.093
26	0.005	-0.006	0.002**	-0.001	-0.003	-0.007	0.940***	-0.037
27	0.011*	-0.006	0.003**	-0.001	-0.031***	-0.005	0.907***	-0.046
28	0.036***	-0.007	0.003	-0.004	-0.051***	-0.013	1.033***	-0.11
29	0.023***	-0.006	0.002	-0.003	-0.036***	-0.007	0.742***	-0.098
30	0.018***	-0.006	0.004	-0.003	-0.040***	-0.007	0.549***	-0.076
31	0.039***	-0.008	0.008	-0.006	-0.034**	-0.015	0.793***	-0.096
32	0.030***	-0.007	0.005	-0.006	-0.048*	-0.025	1.607***	-0.193
33	0.014**	-0.007	-0.001	-0.007	-0.091	-0.066	1.761***	-0.131
34	0.022***	-0.006	0.005	-0.004	-0.038***	-0.01	1.107***	-0.081
35	0.006	-0.011	0.01	-0.012	0.004	-0.027	1.359***	-0.14

* See Appendix 1 for technology codes and corresponding names.

** Country and time fixed effects. We control for regional and technology variables likewise other models in this paper.