Papers in Evolutionary Economic Geography

21.17

The emergence of Artificial Intelligence in European regions: the role of a local ICT base

Jing Xiao & Ron Boschma

Utrecht University Human Geography and Planning

The emergence of Artificial Intelligence in European regions: the role of a local ICT base

Jing Xiao*, Ron Boschma**

May 25, 2021

Abstract:
The purpose of this study is to investigate how a regional knowledge base in Information and Communication Technologies (ICTs) influences the emergence of AI technologies in European regions. Replying on patent data and studying the knowledge production of AI technologies in 233 European regions in the period of 1994 to 2017, our study reveals three results. First, ICTs are a major knowledge source of AI technologies and their importance has been increasing over time. Second, a regional knowledge base in ICTs is highly relevant for regions to engage in AI inventing. Third, the effects of regional knowledge base of ICTs are stronger for regions that recently caught up in AI inventing. Our findings suggest that ICTs play a critically enabling role for regions to diversify into AI technologies, especially in catching-up regions.

Key words: Artificial intelligence (AI), regional diversification, Information and Communications Technologies (ICTs), technological relatedness, General Purpose Technologies (GPTs), Europe

JEL classification: O33, R11, O31

* Centre for Innovation Research (CIRCLE), Department of Design Sciences, Lund University, PO Box 117, 22100 Lund, Sweden. Email: jing.xiao@circle.lu.se.

** Department of Human Geography and Planning, Utrecht University, PO Box 80 115, TC Utrecht, NL 3508, The Netherlands; Stavanger Centre for Innovation Research, UiS Business School, University of Stavanger, Stavanger, N 4036, Norway. Email: r.a.boschma@uu.nl.

1. Introduction

Artificial Intelligence (AI) has been drawing increasing attention in both academic and policy circles, due to its disruptive nature and enormous growth potential (Agrawal, Gans & Goldfarb, 2019; Buarque, Davies, Hynes & Kogler 2020; European Commission, 2018). AI, along with other emerging technologies, are believed to be the core of innovations for the next generation. The diffusion of AI entails new opportunities for a region to expand its technological portfolio and create new growth paths, which matters for the region's structural change and long-term sustainable development. Particularly, for the past several years, AI has been growing rapidly in research and patenting along its technological breakthroughs (WIPO, 2019). How to catch up the rising opportunities from the recent development of AI to preserve EU's technological competence is high on the agenda for the European regions (European Commission, 2020).

What drives the emergence of new growth paths in a region? It has been one of the core topics in the field of Evolutionary Economic Geography. This strand of literature approaches diversification as a process of regional branching: new technologies or activities are more likely to emerge in a region when they are related to the pre-existing local capabilities (Frenken & Boschma, 2007; Boschma & Frenken, 2011b; Boschma, 2017). Relatedness is argued to capture cognitive proximity, which, along with other dimensions, such as geographical or institutional proximity (Boschma, 2005), could facilitate knowledge diffusion within regions and thus explain why related activities are more likely to emerge (Boschma, 2017). This strand of research has often focused on the average effects of relatedness. However, relatedness effects may differ by different types of pre-existing technologies. Technological evolution is argued to be driven by a few radical technologies, termed as "General Purpose Technologies" (GPTs) by Bresnahan and Trajtenberg (1995) to highlight their far-reaching effects that lead to not only technological progress but also substantial changes in economic and social structures. Regions differ substantially in terms of technological and industrial structures as a consequence of previous GPTs, which thus sets the limitations for the emergence of future technologies.

Surprisingly less attention has been paid to how regional branching is influenced by GPTs. In addition to technological relatedness, GPTs has been emphasized as a key tool for smart specialisation policy as the diffusion of GPTs is believed to create new opportunities by co-invention of applications (Foray, David, & Hall, 2009; Montresor & Quatraro, 2017). However, our knowledge of how technological relatedness and GPTs could jointly influence regional technological evolution is limited. Particularly over recent decades, a new wave of disruptive technologies has been emerging and bringing in new transformative pressures for regions to strive for new diversifications. How do the currently predominated GPTs – Information and Communication Technologies (ICTs) play a role in the emergence of next generation of radical technologies? Does this role change along with the diffusion of ICTs? To the best of our knowledge, there has been no relevant study so far.

To fill the gap, this study aims to investigate how regional knowledge base of ICTs influences the emergence of AI technologies in European regions. We argue that ICTs, as the currently predominated GPTs, should play a critical role in breeding the next generation of digital technologies in general and AI technologies in particular. First, ICTs provide knowledge base and building blocks that equip regions with digital capabilities and infrastructures to underpin the local capabilities of capturing AI opportunities. Second, the diffusion of ICTs unlocks new technological opportunities for AI and thus increases recombination possibilities for regional technological diversification.

Recent empirical studies have directed attention to regional diversification processes of newly emerging technologies, such as fuel cell technologies, nanotechnologies, biotechnologies and Industry 4.0 technologies (including AI) (Balland & Boschma, 2021; Colombelli, Krafft & Quatraro, 2014; Feldman, Kogler, & Rigby, 2015; Heimeriks & Boschma, 2014; Lafffi & Boschma 2021; Montresor & Quatraro, 2017; Tanner, 2016). Few studies, however, have examined the regional evolution of AI. One of the main reasons is attributed to the lack of appropriate data (Buarque et al., 2020). Over the last couple of years, EPO (2017) and WIPO (2020) have separately released methods to identify AI patents based on key phrase or patent classification code searching. Among the limited studies on regional development related to AI, one is by Buarque et al. (2020), which focuses on the geographical mapping of AI technologies in European regions and explores the role of AI in regional knowledge network. They find that AI successful regions are more likely to be the regions where AI technologies are most embedded in their knowledge space. Another one is by Balland and Boschma (2021), which focuses on the regional knowledge production of Industry 4.0 technologies (including AI) in general. They find that a new Industry 4.0 technology is more likely to emerge in a European region if the existing technologies in the region are highly related to Industry 4.0 technologies. A very recent study is by Laffi and Boschma (2021), which provides more direct evidence showing that the probability of emergence of Industry 4.0 technologies is higher for regions which specialize in Industry 3.0 technologies. These studies concentrate either on the current position of AI technologies in the knowledge space or on the relationship between Industry 3.0 and Industry 4.0 technologies in general.

The role of GPTs in technological diversification has been neglected in the extant literature. One exception is the study by Montresor and Quatraro (2017). They examine the effects of GPTs by focusing on a group of new generation key enabling technologies, such as industrial biotechnology and nanotechnology. However, there has been no direct evidence exploring how GPTs influence the emergence of AI at regional level. Particularly, to our best knowledge, there has been no studies so far that have explored explicitly which technologies serve as the main knowledge sources of AI technologies.

To explore how a regional knowledge base in ICTs influences the emergence of AI technologies, we build a dataset for the period of 1994 to 2017 based on the patent data from the OECD REGPAT database. To analyze the knowledge source of AI technologies, we conduct a citation analysis to identify the technological fields of the patents that were cited by AI patent applications. We find that instruments and ICTs are two major knowledge sources cited by AI patent applications. Among others, the importance of ICTs, particularly those advanced digital technologies, has become increasingly important over time. In the period of 2012 to 2017, ICTs have surpassed instruments and become the largest knowledge source cited by AI patent applications. In addition, we calculate the average technological relatedness of ICTs to a region's existing knowledge base and model its effects on regional knowledge production of AI. Based on a fixed effects negative binomial model, we find that regions with a high level of technological relatedness of ICTs increases AI inventing. The effects of technological relatedness of ICTs are stronger for regions that recently caught up in AI inventing.

The rest of the paper proceeds as follows. Section 2 reviews briefly the relevant literature and discusses the theoretical background. Section 3 describes the data and method. Section 4 displays the analyses and the findings, and the final section concludes and discusses the paper.

2. Theoretical background

2.1 Relatedness, recombination, and regional diversification

Regional diversification concerns the emergence of new economic activities and new growth paths (Neffke, Henning & Boschma 2011; Isaksen, 2015; Tanner, 2016), which matters for structural change and long-term sustainable development for a region (Content and Frenken, 2016; McCann, 2013). The extant literature has emphasized two critical mechanisms of regional diversification. The first mechanism is recombination process. In his seminal work, Schumpeter (1934) proposes “the carrying out of new combinations”, which is believed as the critical action behind knowledge creation and innovation (Weitzman, 1998), and highlighted as a key mechanism behind regional diversification, enabling recombining and modifying existing capabilities (Boschma & Frenken, 2011b). The other mechanism is knowledge diffusion. Relatedness captures cognitive dimension of proximity, which, along with other dimensions, such as geographical or institutional proximity (Boschma, 2005), are believed to facilitate knowledge transmission and spillovers through promoting interactive learning (Boschma, 2017). Learning allows agents to acquire new knowledge developed by others and improve the chances to create new knowledge through recombination. A burgeoning literature has provided widely supportive evidence for the relatedness hypothesis, no matter when diversification is measured by entry of new products, technologies or industries, or analyzed in different geographical units (such as country, region or city) (see. e.g., Boschma & Frenken, 2011a; Boschma, Balland & Kogler 2015; Boschma, Minondo & Navarro, 2013; Hidalgo, Klinger & Hausmann, 2007; Neffke et al., 2011; Rigby, 2015).

2.2 The role of General Purpose Technologies

The two mechanisms mentioned above suggest that the previous generation of technologies, especially those which are labeled as GPTs (General Purpose Technologies), given their two unique properties, is supposed to play a critical role in breeding the emergence of new generation of technologies. First, GPTs are pervasive in nature, which means GPTs will be applied in a wide range of sectors and eventually penetrate every part of the economy (Bresnahan & Trajtenberg, 1995). The far-reaching effects of GPTs are not only beyond technological progress but also on changes of the behavioral pattern of the entire economy, leading to shifts of “techno-economic paradigm" (Freeman and Perez, 1988; Perez, 2002). GPTs are relatively few, but represent the exemplary technologies behind industrial revolutions, such as steam engine of the first industrial revolution, electricity of the second and ICTs of the third (Helpman & Trajtenberg, 1996). Second, GPTs will unlock complementary innovation opportunities for other sectors along their diffusion processes (Bresnahan, 2012). Technologies are interdependent and cumulative in nature, which means the availability of complementary technologies is a prerequisite that technologies could function or generate economic impacts (Rosenberg, 1979). In some cases, recombination possibilities will only appear after a new technology is invented (Rosenberg, 1979). As a GPT diffuses, it allows the actors in other sectors to recombine their existing technologies with the GPT and create new applications, leading to reconfiguration and evolution of the user sectors' technological portfolios. For example, Fai and Von Tunzelmann (2001) study the evolution of technological scale and scope by following the 32 largest inventing firms over 60 years and find that technological diversity is positively related to the emergence of new technological paradigms. Mendonça (2006; 2009) finds that the emergence of ICT paradigm is a distinct force that drives traditional sectors to diversify into ICTs, not only as users but also as active knowledge producers.

2.3 Diffusion of ICTs as digital base of AI

Building upon the two premises of GPTs, we argue that diffusion of ICTs serves as the digital base for the emergence and development of AI. First, the diffusion of ICTs provides pervasive digital infrastructures for adoption of AI. As indicated by the model of Perez (2002), in the early stage of the diffusion of GPTs, which she refers to as "installation period", GPTs emerge as disruptive technologies and start to reshape the whole economic system by directing to new investment opportunities. One consequence of this period is that a large scale of new infrastructures is set up for future exploitation of the GPTs more efficiently. It is not only that technologies build upon each other, but also on technological paradigms. In the example of the adoption of ICTs, studies show that the energy infrastructure (electricity) plays a fundamental role, especially for low-income countries which face energy constraints (Aebischer & Hilty, 2015; Armey & Hosman, 2016). Apparently, due to the differentiated economic paths and industrial structures, regions differ substantially in terms of the infrastructures related to the previous technological paradigm, which will restrict future diversification possibilities. This "lock-in" effect may especially apply to less developed regions, which have limited recombination opportunities or are in the periphery of knowledge space. E-skills or e-competencies, or the general quality of human capital in a broader sense, constitutes another important dimension of local capabilities that matter for the emergence of new technologies. Castellacci, Consoli and Santoalha (2020) shows that e-skills, measured as the regional intensity of users or developers of ICTs, play a stronger role in regional technological diversification for less developed regions or regions with low levels of relatedness. Similarly, Cohen and Levinthal (1989; 1990) emphasize the importance of previous knowledge in understanding, assimilating and utilizing external knowledge in innovation. This type of "absorptive capacity" is critical for the adoption of new knowledge. For example, based on survey data, a McKinsey report shows that about 75% of AI adopters (firms) reply on their existing digital knowledge and capabilities (Bughin & Van Zeebroeck, 2018).

Second, the diffusion of ICTs unlocks new technological opportunities for AI and thus increases recombination possibilities for regional technological diversification. In the diffusion process, ICTs sectors are evolving with a high speed of updates and iterations of technologies. Several technological shifts, for example, from computer-dominant to internet/web-services dominant technologies, were observed in the past decades. The technological updates provide new opportunities for the users or downstream industries to create new innovational complementarities (Bresnahan & Trajtenberg, 1995). For example, the rise of e-commerce or e-advertisement exhibits a deep penetration of ICTs to traditional sectors, like retailing and advertising industries. Meanwhile, the rapid evolution of ICTs opens up new technological opportunities, acting as key enablers for the advancement of AI. The recent study by Montresor and Quatraro (2017) shows that a new generation of technologies plays a critical role in promoting regional technological diversification in European regions. They highlight that key enabling technologies not only augment the diversity of recombinant technologies but also unlock the recombination constraints. The recent upsurge of AI, to a large extent, benefits from the advances in machine learning, which in turn, depends crucially on the increasing computing power, high-speed connectivity, and availability of large volume of data (WIPO, 2019). In this sense, ICTs can be regarded as a critical external knowledge source for AI, not only feeding new technologies but bridging possibilities for recombination.

3. Data and methodology

3.1 Data

The data we use are from the OECD REGPAT database (January 2020 edition[1]). The OECD REGPAT data have been geocoded and linked to regions across OECD and European countries (see, Maraut, Dernis, Webb, Spiezia & Guellec, 2008, for more details). This gives us a unique opportunity to compare regional differences in terms of knowledge production of AI technologies. In addition, we use patent citation data from OECD (July 2020 edition) to identify the possible knowledge sources of AI patents by tracing the citation flows.

The OECD REGPAT database comprises two types of datasets: patent applications to the EPO (European Patent Office) and patent applications under the Patent Co-operation Treaty (PCT). The PCT is an international patent system, which facilitates applicants to seek for international patent protection. By relying on the patent data under the PCT, we could capture the patent applications with higher technical values because an international patent application usually involves much higher costs. This may also generate less country-based biases in the analysis of cross-section comparison, considering that PCT is an international patent system (Tanner 2016).

Identifying AI patent applications

WIPO has recently developed and published a PATENTSCOPE Artificial Intelligence Index as a search model for AI patent applications (WIPO 2019, 2020). The index comprises key phrases, IPC (International Patent Classification) and CPC (Cooperative Patent Classification) codes and can be used as key search criteria for capturing AI technologies. We use the CPC codes in our analysis to identify AI patent applications, because classification codes can provide a completer, more precise search compared to key phrases. Particularly, the CPC is an extension of the IPC and a more fined grained scheme[2], which may better capture the AI technologies that are scattered in different technological fields.

Based on the CPC codes listed in the index, we have identified 13,781 unique AI patent applications under the PCT from 1980 to 2017[3], accounting for about 4% of all applications during the period. Figure 1 displays the number of AI applications under the PCT over time. The

[1] The REGPAT database (January 2020 edition) derives from the PATSTAT (Worldwide Statistical Patent Database)'s EP Register (Spring 2020 version). The REGPAT database was not updated in the version of July 2020. Therefore, we use the version of January 2020.

[2] CPC has about 250,000 classification entries while IPC has about 70,000 classification entries.

[3] This analysis focuses on the period of 1980 to 2017. First, it is because no AI patent applications under the PCT are identified before 1980 in the REGPAT database. Second, it is because this version of REGPAT database only covers a small part of patent applications in 2018 and 2019 (priority year).

figure shows that AI patent applications started to increase at a faster pace from 1990s particularly from 2012. The rapid growth is attributed to the breakthroughs in machine learning which benefit from increasing computing power, data availability and connectivity over recent years (WIPO 2019). To exhibit a more intuitive picture of the technological base of AI patent applications, we aggregate the frequency of CPC classes of AI technologies (within AI patent applications) into 4-digit level (subclass level). In Table 1, we display the top 10 technological fields[4] of AI technologies. Table 1 shows that AI technologies concentrate mainly in the CPC class of instruments (3-digit level) such as technologies related to recognition of data, digital data processing and computational models, which are related to basic AI techniques such as machine learning. We can also find the presence of AI technologies in the technological areas where AI is applied into practice, such as computer vision, speech recognition, health and transportation.

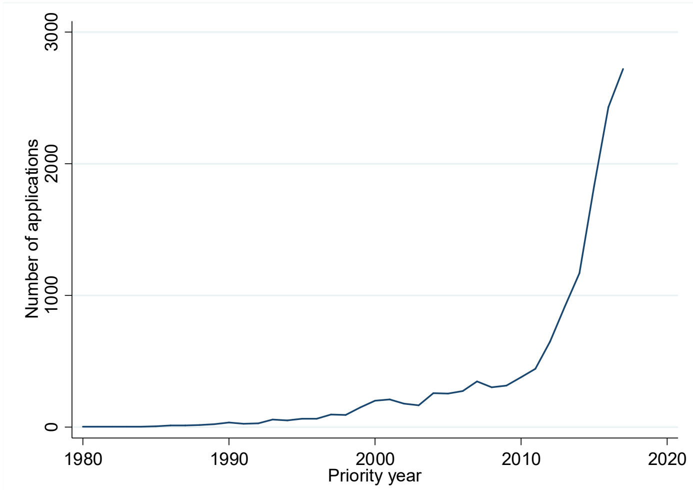

Figure 1. Number of AI patent applications under the PCT over time

[4] We calculate the share of frequency of each 4-digit CPC class within all AI patent applications between 1980 to 2017. The top 10 technological fields account for almost 90% of AI patent applications during the period.

Table 1. Top 10 technological fields (4-digit CPC classes) of AI technologies for the period between 1980 to 2017

CPC	CPC subclass description (4-digit)	CPC class description (3-digit)	Share
G06K	Recognition of data; presentation of data; record carriers; handling record carriers	Instruments	14.23%
A61B	Diagnosis; surgery; identification	Health; Amusement	14.14%
G06F	Electric digital data processing	Instruments	11.86%
G06N	Computer systems based on specific computational models	Instruments	11.41%
G06T	Image data processing or generation, in general	Instruments	10.85%
G05D	Systems for controlling or regulating non-electric variables	Instruments	8.08%
B60W	Conjoint control of vehicle sub-units of different type or different function; control systems specially adapted for hybrid vehicles; road vehicle drive control systems for purposes not related to the control of a particular sub-unit	Transporting	6.48%
G10L	Speech analysis or synthesis; speech recognition; speech or voice processing; speech or audio coding or decoding	Instruments	5.50%
B62D	Motor vehicles; trailers	Transporting	2.75%
G05B	Control or regulating systems in general; functional elements of such systems; monitoring or testing arrangements for such systems or elements	Instruments	2.55%

Source of CPC description: EPO

To assign the patents into regions, we can depend on the location information of either applicants or inventors. However, big firms (as one major group of patent applicants) usually register their patents under their headquarters (Maraut et al., 2008). Using the location information of applicants may therefore bias the geographical patterns of AI inventing activities. Thus, we use the location information of inventors to assign patents into regions. Since our study aims to identify and measure knowledge production/distribution based on the frequency of patent applications instead of assessing the relative regional contribution of inventors from different regions, a non-fractional count of inventors is preferred when we assign patents to regions. We use this non-fractional count as a measure of AI inventing for each region, which means patent applications are counted every time for a region when an inventor is geolocated in this region. Only 8.5% of the 13,781 AI patent applications involve one inventor. About 90% involve 2-10 inventors. In terms of the geographical distribution of AI inventing, the top 20 countries account for over 96% of all AI patent applications under the PCT from 1980 to 2017, including the US, Japan, China, Germany, South Korea, the UK, the Netherlands, Canada, France, Israel, Sweden, Australia, India, Switzerland, Spain, Singapore, Finland, Italy, Ireland and Denmark.

AI inventing in Europe

Since our main interest is on AI inventing in European regions, our analysis only includes the regions within EU27+3 countries[5]. As AI technologies are still in their early stage of development, our analysis starts from 1994 when AI technologies begun to develop and diffuse at a faster pace (WIPO, 2019). This ends up with 233 European regions (NUTS2 level) with AI

[5] EU 27+3 countries include Austria, Belgium, Bulgaria, Bulgaria, Croatia, Cyprus, Czechia, Denmark, Estonia, Portugal, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Poland, Romania, Slovakia, Slovenia, Spain, Sweden, plus the UK, Switzerland, and Norway.

inventing for the period of 1994 to 2017[6]. Figure 2 displays the histogram of the number of AI patent applications[7]. About 97% of the regions have less than 10 AI patent applications during the whole period. The distribution of AI inventing is highly skewed over time and space.

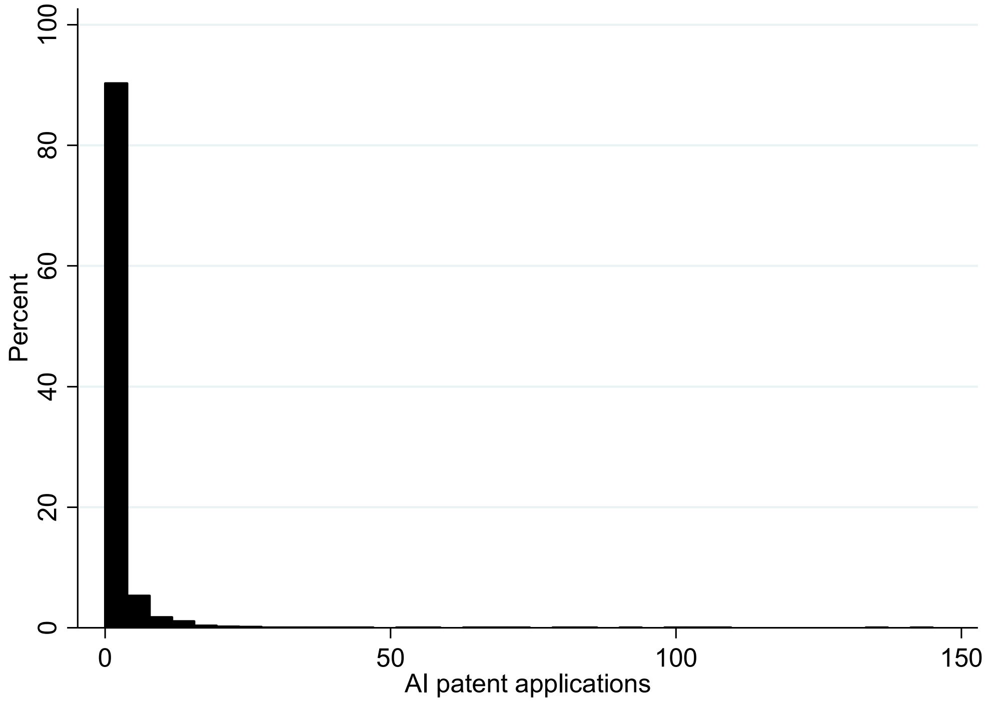

Figure 2. Histogram of the number of AI patent applications of European regions

Figure 3 maps the number of AI patent applications across the European regions over four periods: 1994-1999, 2000-2005, 2006-2011, 2012-2017, respectively. During the early period, there are only a limited number of regions with AI patent applications and AI inventing concentrates in three German regions, namely Stuttgart (DE11), Oberbayern (DE21) and Mittelfranken (DE25). Over time, it is noted that more regions are involved in AI patent applications. Especially, during the period of 2012 to 2017, 53 regions are found to have more than 20 AI patent applications. To exhibit the hotspots of AI inventing, we map the share of AI patent applications of European regions over the same periods, see Figure 4. It is noted that the hotspots of AI patent applications concentrate in a few regions in Western European countries, such as Germany, the Netherlands, and France. From the period of 2012-2017, the top 5 regions account for almost 40% of all AI patent applications, including North Brabant (NL41) in the Netherlands, Oberbayern (DE21) and Stuttgart (DE11) in Germany, Inner London (UKI1) in the UK, and Ile de France (FR10) in France.

---

[6] Some inventors are only assigned to a country but not an accurate region, we removed these inventors (61 inventors in 13 countries). The regions with no AI patent application in the whole period are not included as we will use fixed-effects estimator in the econometric analysis.

[7] The observation is on a yearly basis.

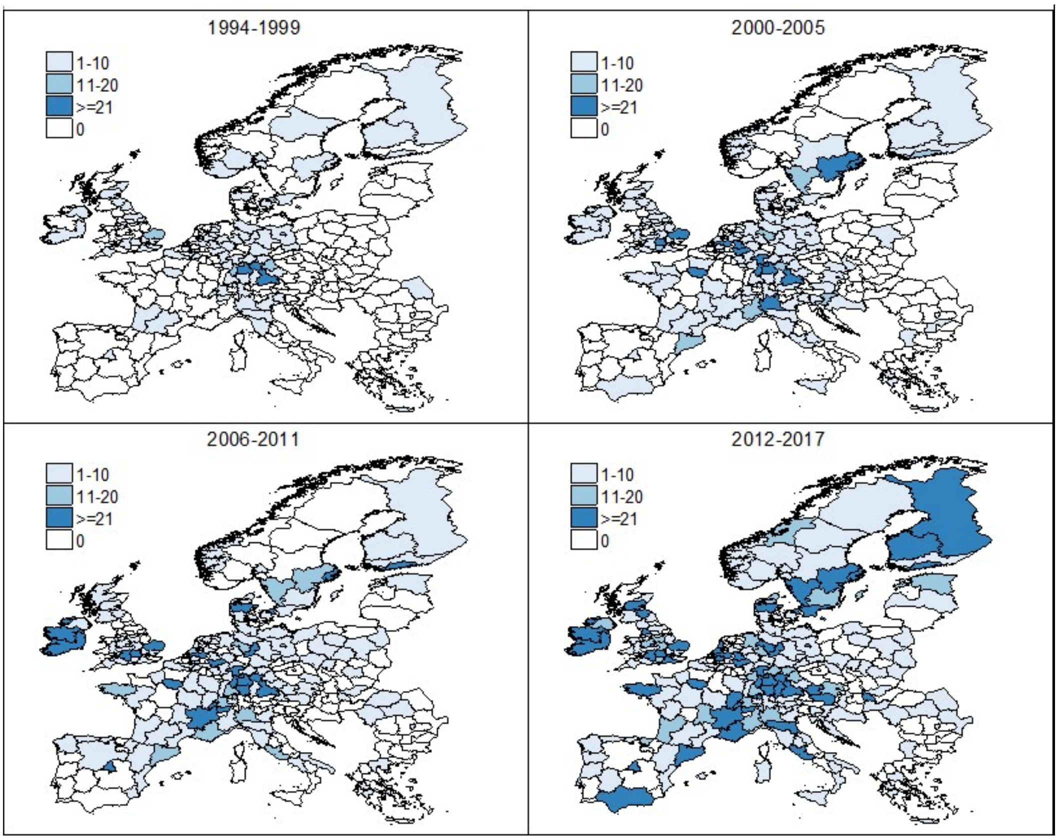

Figure 3. Number of AI patent applications of European regions

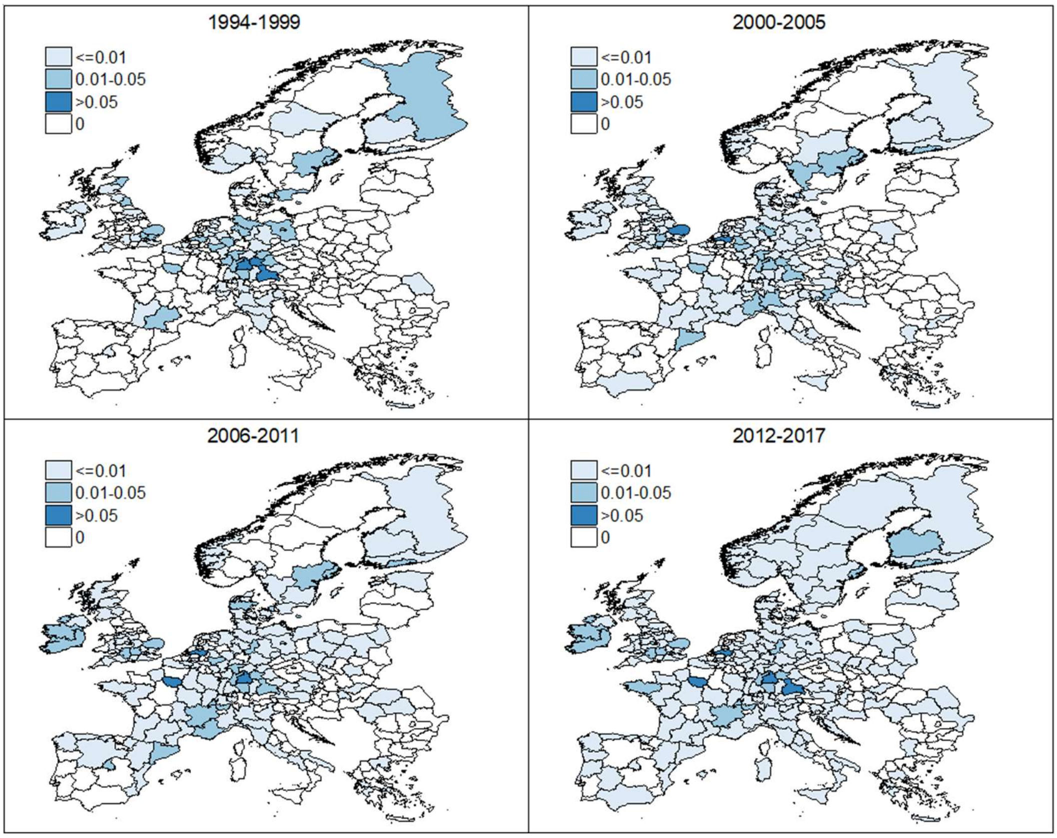

Figure 4. Share of AI patent applications of European regions

3.2 Variables for econometric analysis

Dependent variable: measuring knowledge production of AI technologies

One aim of this study is to examine the role of technological relatedness of ICTs in the emergence of AI technologies. We use the number of AI patent applications as the indicator of production of AI technologies in a region. To avoid the situation where some regions may have a very small number of counts, we divide the whole analysis period into eight sub-periods[8] and use the sum of each sub-period of each region as the dependent variable. In this paper we model regional number of AI technologies as a function of technological relatedness of ICTs. Compared to an entry model in which the probability of a new technological specialization is modeled as a function of the relatedness of focal technology to the local structure of existing technologies, we think our model may better capture the regional variation in AI knowledge production and how it is related to local ICT base at the early development stage of AI.

Independent variables: measuring relatedness with existing technologies in a region

To indicate technological relatedness of ICTs to a region's existing knowledge base, we develop a variable measured as the average density of relatedness of ICTs to a region's existing knowledge base. The variable is developed in two steps. We first calculate the proximity between all technologies. To this end, we conduct a co-occurrence analysis to measure the relatedness between technologies. This approach is developed by Hidalgo et al. (2007) where they measure proximity of products based on the likelihood of simultaneous occurrence of two exported products in a country, given the assumption that related products share similar factor endowments or capabilities. This approach has been widely adopted in previous studies on industrial diversification or regional branching (see, e.g., Cortinovis et al., 2017; Xiao et al., 2018; Boschma et al., 2013; Hausmann et al. 2010). A patent usually involves multiple classification codes to indicate the technological fields covered by the patent. We assume that all the co-classified technological fields share technological relatedness and their proximity to each other can be captured by the likelihood of their co-occurrence in a patent. We calculate the proximity among all technological fields (at 4-digit IPC level[9]) by the likelihood of their co-occurrence in non-AI patent applications[10], as shown in Equation (1).

$$\varphi_{i,j,t} = \min\{P(x_{i,t}|x_{j,t}), P(x_{j,t}|x_{i,t})\} \tag{1}$$

where φ indicates the proximity index. The index is the minimum conditional probability that a patent involves one technological filed given that it involves another technological filed. The second step is to link the proximity index with a region's existing knowledge base. A region's knowledge base is indicated by the collection of technological fields of its patent portfolio. Again, to avoid potential endogeneity, we exclude all AI patent applications when identifying a region's existing knowledge base. The average density of relatedness of ICTs to a region's existing knowledge base is calculated as shown in Equation (2).

---

[8] The eight sub-periods include 1994-1996, 1997-1999, 2000-2002, 2003-2005, 2006-2008, 2009-2011, 2012-2014, 2015-2017.

[9] IPC is used here because most existing definitions of ICT are based on IPC classification.

[10] We exclude AI patent applications to avoid potential endogeneity biases.

$$\overline{d_{\iota,r,t}} = \left(\frac{\sum_k \varphi_{i,k,t}\, x_{k,r,t}}{\sum_k \varphi_{i,k,t}}\right) \quad (2)$$

where the subscript *i* or k refers to a technological filed; $x_{k,r,t}$ is a dummy variable to show whether technology *k* is present in region *r* at year *t*. $d_{i,r,t}$ is the density around technology *i* in region *r* at year *t*, calculated as the sum of proximities of technology *i* to all technologies that are present in region *r* at year *t* divided by the sum of proximities of technology *i* to all technologies. The density varies between 0 to 1. A higher density means a higher level of relatedness of technology *i* to the technologies that are present in region r. At the end, we take the average density of all ICTs for each region. We use a broad definition of ICTs to calculate the average density of relatedness of ICTs to a region's existing knowledge base. The definition of ICTs is elaborated in Section 4.1 at below.

4. Analyses and results

4.1 Knowledge sources of AI patents

As a newly emerging technology, the development of AI may draw upon various sources of established knowledge. To identify the knowledge sources of AI patents, we rely on the citation analysis, where the patent citation data are used as a proxy to measure knowledge flows or spillovers (Jaffe & Trajtenberg, 1998; Jaffe, Trajtenberg & Fogarty, 2000). To quantify the relative intensity of knowledge flows among different sources, we use the technology classification to group the technological fields of the patents that were cited by the AI patents.

The classification we use is mainly based on the typology by Schmoch (2008, updated in 2019) which groups all patentable technological fields into five general categories (based on IPC classification), including electrical engineering, instruments, chemistry, mechanical engineering, and other fields. *Electrical engineering* further constitutes 8 sub-categories, including electrical machinery, apparatus, energy; audio-visual technology; telecommunications; digital communication; basic communication processes; computer technology; IT methods for management; and semiconductors. To capture the role of ICTs in a broader sense and the ICTs at the technological frontier respectively, we use two definitions of ICTs in the citation analysis: the broad definition (defined as the category of electrical engineering excluding the sub-category of electrical machinery, apparatus, energy) and the restrictive definition (defined as those ICTs that are categorized into high technology by Eurostat (2006), including computer and automated business equipment, semiconductors, and communication technology). Accordingly, we revise Schmoch's typology into 6 general categories in our analysis:

- electrical machinery, apparatus, energy,
- ICTs (broad definition)
- instruments,
- chemistry,
- mechanical engineering,
- other fields.

More than 88% of the IPC classes defined by restrictive definition of ICTs fall within the broad definition of ICTs. The rest concentrate in the IPC class of "B41J" which is grouped into the category of mechanical engineering in Schmoch's typology. In our analysis with AI patent applications, the restrictive definition is a subset of the broad definition because no cited technological filed falls within “B41J”.

To measure the intensity of knowledge flows between technological categories and AI, we calculate the share of citations: the cited number of each technological category divided by the total cited number. We report the results over four periods, 1994-1999, 2000-2005, 2006-2011, 2012-2017, respectively in Table 2.[11] The two major knowledge sources of AI patent applications are instruments and ICTs (broad definition). For AI patent applications between 1994 to 1999, the share of cited number is 46% for instruments and 35% for ICTs (broad definition) respectively. The share of ICTs (restrictive definition) is 17%. Over time, the relative importance of ICTs (measured by both broad definition and restrictive definition) is increasing. At the same time, the share of instruments is decreasing. For AI patent applications between 2012 to 2017, ICTs have become the largest knowledge source of AI patent applications. About 43% of cited technologies are from the category of ICTs (broad definition). During the same period, the share of ICTs (restrictive definition) has increased to 25%. This indicates an increasing importance of digital base for the recent development of AI, which is consistent with the recent trend in AI patenting where machine learning has been predominating in the patent applications related to AI techniques and AI related patents (WIPO, 2019).

[11] To be able to compare the citation patterns between periods, we restrict to the cited patents that were published in the recent 11 years for each period. For example, for period of 1994 to 1999, we include cited patents there were published between 1989 to 1999. This analysis includes all AI patents identified in the PCT database.

Table 2. The share of cited number of each technological category by AI patent applications

	Electrical machinery, apparatus, energy	ICT		Instruments	Chemistry	Mechanical engineering	Other fields
		Broad	Restrictive				
				1994-1999			
Cited number	4	73	35	94	3	21	11
Share	2%	35%	17%	46%	1%	10%	5%
				2000-2005			
Cited number	24	222	94	232	17	41	10
Share	4%	41%	17%	42%	3%	8%	2%
				2006-2011			
Cited number	46	350	205	391	21	116	23
Share	5%	37%	22%	41%	2%	12%	2%
				2012-2017			
Cited number	59	1434	832	1195	86	457	71
Share	2%	43%	25%	36%	3%	14%	2%

4.2 The effects of technological relatedness of ICTs

We conduct an econometric analysis to test the effects of technological relatedness of ICTs to a region's existing knowledge base on regional knowledge production of AI inventing. The final dataset for the econometric analysis is a balanced panel covering 233 European regions over 8 periods. Table 3 display a summary of variable description and descriptive statistics. Table A1 in the Appendix shows the correlation matrix between the variables. In addition to the dependent and independent variables, we include a control variable of population to account for regional differences in size that are changing over time. Since the dependent variable is the number of AI applications, we expect a positive relationship between population size and regional number of AI patent applications. The data of population are from Eurostat. Population has missing values because of the changes of NUTS classification systems over time[12]. Moreover, we include the dummy variable for time periods to control for time effects in general. The inclusion of the period dummy variables is expected to capture the major time-varying heterogeneity between regions. Independent and control variables are measured at one year before each period. In our dataset, there are about 53% of observations have no AI inventing and the distribution of AI inventing is highly skewed.

Table 3. Variable description and summary statistics

Variables	Description	Obs	Mean	Median	Std. Dev.	Min.	Max.
AI_inventing	Number of AI patent applications	1,864	4.34	0	17.47	0	335
Ave_density	Average density of technological relatedness	1,864	0.28	0.24	0.23	0	0.94
Pop	Population	1,661	14.27	14.27	0.68	12.40	16.30

Note: There are missing values for the variable of population, which is because of the changes of NUTS classification systems over time.

Since the dependent variable is the number of AI patent applications, we use count model to model the effects of technological relatedness on AI inventing. Fixed effects are used to account for unobserved heterogeneity that is constant over time at regions. Since our data suffer from the problem of overdispersion (variance is higher than mean), this suggests a negative binomial model. However, many studies indicate that the method for (conditional) fixed effects negative binomial regression that many statistical software products (such as Stata) depend on is not valid because it fails to control for unchanging covariates (Allison & Waterman 2002; Greene 2005; Guimarães, 2008). Following the suggestion by Allison (2012), we use the unconditional fixed effects negative binomial model by including dummy variables for each region as our benchmark estimation model.

To facilitate interpretation, we standardize technological relatedness and population in the regressions. The results are reported in Table 4. In Specification (1), we only include the main predictor. In Specification (2), we add population as control variable. In Specification (3), we include period dummies as control variables.

[12] The changes of NUTS classification systems over time make it difficult to consistently trace the regional statistics over a long period of time for the regions which are affected. This is one reasons why we do not include more regional-level control variables as there will be many missing values. Another reason is because we include period dummies to control for time effects, which we believe would capture the major time-varying heterogeneity.

Table 4. The effects of technological relatedness of ICTs on AI inventing

Variables	(1)	(2)	(3)
Ave_density	1.618***	1.242***	0.479***
	(0.0923)	(0.101)	(0.110)
Pop (log)		5.630***	
		(0.667)	
Constant	-0.356	-0.209	-1.211***
	(0.477)	(0.450)	(0.429)
Obs	1,864	1,661	1,864
Period dummies	No	No	Yes
Log likelihood	-3081.0321	-2723.1711	-2884.9414
LR chi2	1491.38	1424.67	1883.56
Prob > chi2	0.0000	0.0000	0.0000

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The coefficient of negative binomial model is interpreted as the expected difference in the logs of expected counts of the dependent variable given one-unit change of the independent variable, while all the other variables are held constant. From Specification (1), we find that if the average density of technological relatedness of ICTs was to increase by one-unit, the expected difference in the logs of expected counts of AI patent applications would increase by 1.618 unit, while all the other variables are held constant. The significant coefficient reveals a positive effect of average density of technological relatedness of ICTs on regional AI inventing. In Specification (2) when we include the control variable of population, the positive relationship between technological relatedness and regional AI inventing is still significant, although the magnitude decreases slightly. For the control variable, as expected we find a significantly positive effect of population on AI inventing. In Specification (3) when we include period dummies as control variables, the positive coefficient of technological relatedness is still statistically significant, but the magnitude decreases by about two thirds. The results show that technological relatedness of ICTs to a region's existing knowledge base is an important predictor for AI inventing at European regions. But the effect of technological relatedness on AI inventing is reduced when time effects are accounted for.

Compared to the other global players, such as the US and China, Europe has been lagged behind in the investments of the first waves of AI and related technologies (European Commission, 2018; WIPO, 2019). As shown in Figure 3 in Section 3.1, most AI inventing concentrates in a few German regions in the early period. However, the diffusion of AI technologies has been accelerating recently, especially since 2012. Many European regions have been catching up, reflected by the increase of AI patent applications and the spread of AI inventing to more European regions. To examine whether there are any catch-up effects of technological relatedness of ICTs, we create a dummy variable (*catchup*) with 1 indicating the regions which transit from without AI patent application in the early period (1994-2005) to with AI patent applications in the recent period (2006-2017). We include an interaction term between technological relatedness and *catchup* to indicate the catch-up effects. To test the robustness of the results, we use different thresholds to determine whether it is a catch-up region. The results are reported in Table 5.

Table 5. The catch-up effects of technological relatedness of ICTs on AI inventing

Variables	≥5	≥10	≥15
Ave_density	0.328***	0.391***	0.419***
	(0.113)	(0.111)	(0.110)
Ave_density*catchup	2.050***	2.409***	2.897***
	(0.380)	(0.550)	(0.740)
Constant	-1.173***	-1.182***	-1.187***
	(0.423)	(0.424)	(0.425)
Obs	1,864	1,864	1,864
Period dummies	Yes	Yes	Yes
Log likelihood	-2865.9461	-2871.3893	-2873.513
LR chi2	1921.55	1910.67	1906.42
Prob > chi2	0.0000	0.0000	0.0000

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

As reported in Table 5, the coefficient of *Ave_density*catchup* indicates the catch-up effect of technological relatedness of ICTs on AI inventing. The first column displays the results when the catch-up regions are identified if they have no AI patent applications in the period of 1994 to 2005 but have at least 5 AI patent applications in the period of 2006 to 2017. The significant coefficient of the interaction term shows a strong positive catch-up effect of technological relatedness of ICTs. The effect of technological relatedness on AI inventing decreases slightly but still significant. In the second and third column, we test the results by using different thresholds of defining catch-up regions. The coefficients of interaction terms are both significantly positive and the magnitude increases as the threshold increases. The results show that ICTs are an important enabler for regions which caught up in AI inventing. The catch-up effects are stronger for regions which caught up fast.

5. Robustness check

In the econometric analysis, we use the average density of relatedness of ICTs to a region's existing knowledge base to indicate regional knowledge base of ICTs. To check whether our main findings are sensitive to a different measure of independent variable, we employ an alternative indicator – related variety within ICTs for a robustness check. Related variety is a measure to capture both relatedness and variety across activities in a region. The literature on regional innovation has widely discussed the role of related variety, such as related industries, in providing opportunities for recombination of knowledge and facilitating regional innovation and growth. (Frenken, Van Oort & Verburg, 2007, Neffke et al., 2011, Boschma et al., 2013). Following Frenken et al. (2007), we calculate related variety within ICTs as the weighted sum of entropy at the level of five-digit IPC within each three-digit IPC within ICTs, as shown in Equation (3a) and (3b).

$$RV = \sum_{s=1}^{S} P_s \, H_s \tag{3a}$$

$$H_s = \sum_{i \in s} \frac{P_i}{P_s} \, log_2 \, (\frac{1}{P_i/P_s}) \tag{3b}$$

where the subscript i denotes a five-digit IPC which is exclusively under a three-digit IPC s; P refers to the share of patent applications; H_s refers to the five-digit variety within each three-digit IPC. We re-estimate the benchmark model and the models with interaction term based on different thresholds, respectively, and report the results in Table 6.

From Table 6, it is noted that related variety within ICTs shows a significantly positive effect on AI inventing, even though with a lower magnitude than technological relatedness of ICTs. The catch-up effects are only significant when the threshold is set up at ⩾10. One possible explanation is that the variable of related variety within ICTs tends to capture the general composition of knowledge base within ICTs. By contrast, the variable of regional relatedness of ICTs tends to capture the specific relatedness between ICTs and local knowledge base. Another explanation is that we have many regions with missing values for the variable of related variety within ICTs. These regions have limited number of ICTs and thus no variation in the share of patent applications between 3-digit IPC and 5-digit IPC. The reduced number of observations may lead to the insignificance of results in some specifications. But even with the reduced number of regions, it is noted that the sign of catch-up effect is still positive across the specifications with different thresholds and the magnitude tends to increase as the threshold increases.

Table 6. Robustness check: related variety within ICTs as independent variable

Variables	Benchmark model	Catch-up effects ≥5	Catch-up effects ≥10	Catch-up effects ≥15
RV	0.135**	0.118*	0.102*	0.119**
	(0.0581)	(0.0617)	(0.0601)	(0.0589)
RV*catchup		0.136	0.464**	0.525
		(0.170)	(0.229)	(0.337)
Constant	-1.240***	-1.237***	-1.233***	-1.232***
	(0.417)	(0.418)	(0.418)	(0.418)
Obs	1,591	1,591	1,591	1,591
Period dummies	Yes	Yes	Yes	Yes
Log likelihood	-2748.2313	-2747.9047	-2745.9997	-2746.8949
LR chi2	1670.23	1670.89	1674.70	1672.91
Prob > chi2	0.0000	0.0000	0.0000	0.0000

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Recall that we use the broad definition to define ICTs when calculating the technological relatedness in Section 3.2. This may raise a concern of whether our findings are sensitive to change of definition of ICTs. To address this concern, we use the restrictive definition of ICTs for a robustness check. We re-estimate the effects of technological relatedness of ICTs without the interaction terms, with interaction term based on different thresholds, respectively, and report the results in Table 7. The results show that our main findings hold. When we use the restrictive definition of ICTs, both the effects of technological relatedness and the catch-up effects are a relatively weaker than when ICTs are based on the broad definition. This may indicate that what matters for the emergence and catch-up of AI inventing reside more in the ICTs in a broad sense than those advanced ICTs.

Table 7. Robustness check: based on restrictive definition of ICTs

Variables	Benchmark model	Catch-up effects ≥5	Catch-up effects ≥10	Catch-up effects ≥15
Ave_density	0.381***	0.207*	0.283**	0.321***
	(0.112)	(0.115)	(0.113)	(0.112)
Ave_density*catchup		1.871***	2.176***	2.503***
		(0.337)	(0.487)	(0.657)
Constant	-1.257***	-1.223***	-1.231***	-1.234***
	(0.430)	(0.423)	(0.425)	(0.426)
Obs	1,864	1,864	1,864	1,864
Period dummies	Yes	Yes	Yes	Yes
Log likelihood	-2888.582	-2868.3952	-2874.5569	-2877.939
LR chi2	1876.28	1916.66	1904.33	1897.57
Prob > chi2	0.0000	0.0000	0.0000	0.0000

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

6. Discussion and conclusion

Through the lens of regional technological diversification, this paper focused on two specific research questions. How important are ICTs for emergence of AI technologies? And how does regional knowledge base of ICTs influence knowledge production of AI of European regions? Based on the patent data from the OECD REGPAT database, our findings show that ICTs are a major knowledge source of AI technologies and their importance has been increasing over time. We also find that technological relatedness of ICTs to a region's existing knowledge base is an important predictor for the emergence of AI inventing in European regions. Especially, the effects of technological relatedness of ICTs are stronger for regions which recently caught up in AI inventing. Our findings indicate that local infrastructure and capabilities of ICTs serve as the digital base for the emergence and development of AI in European regions. Meanwhile, the development of ICTs itself also unlocks new technological possibilities. Both effects display the enabling nature of GPTs not only feeding new technologies but bridging possibilities for recombination.

The contribution of this paper is three-fold. First, our study theoretically contributes to the literature on Evolutionary Economic Geography by providing new insights of how regional branching is influenced by the diffusion of GPTs. Although technological relatedness and GPTs have been separately emphasized as key tools for smart specialisation policy (S3) (Boschma & Giannelle, 2014; Foray, David, & Hall, 2009), few studies that have investigated how GPTs influence regional diversification and development through the mechanism of technological relatedness. Montresor and Quatraro (2017) is one exception, which explores the role of GPTs in regional branching. But they focus on GPTs as a group of new generation key enabling technologies. This raises a question of whether the new emerging technologies could capture the two properties of GPTs fully. Our findings show that the role of ICTs may go beyond the advanced technologies but resides more in ICTs in a broader sense. Our findings suggest that future studies could go beyond the few key enabling technologies but have a more holistic view to investigate the successive nature of technological evolution. In addition, unlike previous studies, we used citation analyses to exhibit how important ICTs are as one knowledge source of AI and how the importance changes over time.

Secondly, our study methodological contributes to the literature on regional diversification. In the recent studies that focus on regional diversification processes of newly emerging technologies, relatedness is usually measured as the proximity of focal technologies to the local structure of existing technologies. The proximity between technologies is specified by a "technology space", which is usually developed based on the frequency of co-occurrence of technologies in a specific relation, such as co-location in a region, or co-classification in a patent. This approach is useful when the focal technological are stable and mature. But it may set limitations when it is used to measure relatedness of newly emerging technologies, particularly when they are still in the early stage of development. For example, the definition of AI is still fuzzy and has been updating along the fast development of the field (EPO, 2017, WIPO, 2019). In the case of patents, the patent classification codes, such as IPC and CPC, might not have been updated to take full account of emerging technologies. In this sense, the focus only on the relatedness of AI technologies in the current knowledge network may not capture the full picture of its diversification process, as the proximity may not be stable to capture the full picture between AI and other technologies. In our study, instead of focusing on the role of regional knowledge base of AI, we pay attention to the role of regional knowledge base of ICTs. It is not only a relevant technology for AI inventing but also a mature GPT, which is more stable to capture regional knowledge base. This may provide a new view for those studies which aim to investigate the role of relatedness in regional branching of emerging technologies.

Third, our findings also suggest some policy implications. As discussed above, our findings suggest that future regional policies may consider going beyond advanced enabling technologies but pay attention of the role of GPTs in a broader sense in regional development. In addition, past European regional policies on digital technology and AI have paralleled to each other (European Commission, 2016; 2018). For example, e-infrastructure has been addressed in the policies in promoting EU's digital future (European Commission, 2016) and AI technology separately (European Commission, 2018). Our findings indicate a close and successive relationship between digital technology and AI and thus suggest many initiatives or investment opportunities could be jointly coordinated and designed in the future policies.

The new wave of technological change gives new momentum for the field of Evolutionary Economic Geography. It may not only generate new academic debates in terms of how regions embrace the opportunities and challenges arising from the new technologies but also influence the policy approach to integrate the role of technological change in future policy design. We hope this study could attract more future studies to improve our understanding of the micro foundation of how GPTs influence regional diversification.

References

Aebischer, B., Hilty, L.M. (2015) The Energy Demand of ICT: A Historical Perspective and Current Methodological Challenges. In: L.M. Hilty, B. Aebischer (eds) ICT Innovations for Sustainability, Advances in Intelligent Systems and Computing. vol 310. Springer, Cham. https://doi.org/10.1007/978-3-319-09228-7_4.

Allison, Paul, Richard Waterman (2002), Fixed effects negative binomial regression models. In Ross M. Stolzenberg (ed.), Socological Methodology 2002. Oxford: Basil Blackwell.

Allison, Paul (2012), Beware of Software for Fixed Effects Negative Binomial Regression. Statistical Horizons. URL: https://statisticalhorizons.com/fe-nbreg. Date of access: Oct.25, 2020.

Agrawal, A., Gans, J., Goldfarb, A. (2019). Introduction to "The Economics of Artificial Intelligence: An Agenda". In A. Agrawal, J. Gans, A. Goldfarb (eds). The Economics of Artificial Intelligence: An Agenda. Chicago: The University of Chicago Press.

Armey, L.E., Hosman, L. (2016). The centrality of electricity to ICT use in low-income countries. Telecommunications Policy. Volume 40, Issue 7, p. 617-627. https://doi.org/10.1016/j.telpol.2015.08.005.

Balland, P. and R. Boschma (2021). Mapping the potentials of regions in Europe to contribute to new knowledge production in Industry 4.0 technologies, Regional Studies. DOI: 10.1080/00343404.2021.1900557

Boschma, R. (2005). Proximity and Innovation: A Critical Assessment, Regional Studies, 39:1, 61-74. DOI: 10.1080/0034340052000320887.

Boschma, R. (2017). Relatedness as driver of regional diversification: a research agenda. Regional Studies. 51:3, 351-364. DOI: 10.1080/00343404.2016.1254767.

Boschma, R., Balland, P. A., Kogler, D. F. (2015). 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 24 (1): 223–50. DOI:10.1093/icc/dtu012.

Boschma, R., Frenken, K. (2011a). The emerging empirics of evolutionary economic geography. Journal of Economic Geography. 11, 295–307. https://doi.org/10.1093/jeg/lbq053.

Boschma, R., Frenken, K. (2011b) Technological relatedness, related variety and economic geography. In P. Cooke, B. Asheim, R. Boschma, R. Martin, D. Schwartz, F. Todtling (eds). The Handbook on Regional Innovation and Growth. Chapter 14. Cheltenham: Edward Elgar.

Boschma, R., Giannelle, C. (2014). Regional Branching and Smart Specialisation Policy. S3 Policy Brief Series No. 06/2014. Brussels, Belgium: European Commission. https://s3platform.jrc.ec.europa.eu/documents/20182/114990/JRC88242_Regional_Branching_and_S3Policy.pdf/6984a03c-12c7-439a-a311-8289032f6d20.

Boschma, R., Minondo, A., Navarro, M. (2013) The Emergence of New Industries at the Regional Level in Spain: A Proximity Approach Based on Product Relatedness. Economic Geography 89(1): 29-51. DOI:10.1111/ecge.2013.89.issue-1.

Bresnahan, T. (2012). Generality, Recombination, and Reuse. In J. Lerner, S. Stern (eds) The Rate and Direction of Inventive Activity Revisited. 611-656. Chicago: University of Chicago Press. http://www.nber.org/chapters/c12374.

Bresnahan, T., Trajtenberg, M., (1995). General purpose technologies 'Engines of growth'? Journal of Econometrics. Volume 65, Issue 1, Pages 83-108. https://doi.org/10.1016/0304-4076(94)01598-T.

Buarque, B.S., Davies, R.B., Hynes, R.M., Kogler, D.F. (2020). OK Computer: the creation and integration of AI in Europe. Cambridge Journal of Regions, Economy and Society. Volume 13, Issue 1, March 2020, Pages 175–192. https://doi.org/10.1093/cjres/rsz023.

Bughin, J., Van Zeebroeck, N. (2018). Artificial intelligence: Why a digital base is critical. McKinsey Quarterly. URL: https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/artificial-intelligence-why-a-digital-base-is-critical#.

Castellacci, F., Consoli, D., Santoalha, A. (2020). The role of e-skills in technological diversification in European regions. Regional Studies. 54:8, 1123-1135. DOI: 10.1080/00343404.2019.1681585.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and learning: The two faces of R&D. The Economic Journal, 99(397), 569-596. https://www.jstor.org/stable/2233763.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. Administrative Science Quarterly, 35(1), 128-152. https://www.jstor.org/stable/2393553.

Colombelli, A., Krafft, J., Quatraro, F. (2014). The emergence of new technology-based sectors in European regions: A proximity-based analysis of nanotechnology. Research Policy, 43 (10), 1681–1696. doi:10.1016/j.respol.2014.07.008.

Content, J., Frenken, K. (2016). Related variety and economic development: a literature review. European Planning Studies. 24(12): 2097–2112. https://doi.org/10.1080/09654313.2016.1246517.

Cortinovis, N., Xiao, J., Boschma, R., Van Oort, F. (2017). Quality of government and social capital as drivers of regional diversification in Europe. Journal of Economic Geography. 17(6), 1179-1208. DOI: https://doi.org/10.1093/jeg/lbx001.

EPO. (2017). Patents and the Fourth Industrial Revolution. Munich, Germany: European Patent Office. URL: epo.org/4IR.

European Commission. (2016). European Cloud Initiative - Building a competitive data and knowledge economy in Europe. COM (2016) 178 final. European Commission.

European Commission. (2018). Artificial Intelligence: A European Perspective. Luxembourg: Publications Office of the European Union. DOI: 10.2760/11251.

European Commission. (2020). White Paper on Artificial Intelligence A European approach to excellence and trust. COM (2020) 65 final. European Commission.

Eurostat. (2006). Eurostat indicators on High-tech industry and Knowledge – intensive services. Annex 6 – High-tech aggregation by patents. https://ec.europa.eu/eurostat/cache/metadata/Annexes/htec_esms_an6.pdf.

Fai, F., Von Tunzelmann, N. (2001). Scale and Scope in Technology: Large Firms 1930/1990. Economics of Innovation and New Technology. 10:4, 255-288. DOI: 10.1080/10438590100000011.

Feldman, M.P., Kogler, D.F., Rigby, D.L. (2015). rKnowledge: The Spatial Diffusion and Adoption of rDNA Methods. Regional Studies, 49:5, 798-817. DOI: 10.1080/00343404.2014.980799.

Foray, D., David, P. A., Hall, B. H. (2009). Smart specialisation—The concept. Knowledge Economists Policy Brief. Brussels, Belgium: European Commission. http://ec.europa.eu/in vest-in-research/pdf/download_en/kfg_policy_brief_no9.pdf.

Freeman, C., Perez, C. (1988) Structural crises of adjustment: business cycles and investment behaviour. In G. Dosi, C. Freeman, R. Nelson, L. Soete, G. Silverberg (eds.), Technical Change and Economic Theory. London: Frances Pinter. pp. 38–66.

Frenken, K., Boschma, R. (2007). A theoretical framework for evolutionary economic geography: industrial dynamics and urban growth as a branching process. Journal of Economic Geography. 7(5), September 2007, Pages 635–649. https://doi.org/10.1093/jeg/lbm018.

Frenken, K., Van Oort, F. & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. Regional Studies, 41(5), 685–697. https://doi.org/10.1080/00343400601120296.

Greene, William (2005), Functional form and heterogeneity in models for count data. Foundations and Trends in Econometrics 1: 113–218. DOI: 10.1561/0800000008.

Guimarães, P., (2008), The fixed effects negative binomial model revisited. Economics Letters, 99: 63-66. https://doi.org/10.1016/j.econlet.2007.05.030.

Heimeriks, G., Boschma, R. (2014). The path- and place-dependent nature of scientific knowledge production in biotech 1986–2008. Journal of Economic Geography. 14 (2): 339–364. https://doi.org/10.1093/jeg/lbs052

Helpman, E., Trajtenberg, M. (1996). Diffusion of General Purpose Technologies. National Bureau of Economic Research Working Paper Series. No. 5773. September 1996. DOI 10.3386/w5773.

Hidalgo, C. A., Klinger, B., Barabási, A. L., Hausmann, R. (2007). The Product Space Conditions the Development of Nations. Science. 317(5837), 482-487.

Hausmann, R., Hidalgo, C. A. (2010) Country diversification, product ubiquity, and economic divergence. Working Paper No. 201. Cambridge, Mass.: Center for International Department, Havard University.

Isaksen, A. (2015). Industrial development in thin regions: Trapped in path extension? Journal of Economic Geography 15 (3): 585–600. DOI:10.1093/jeg/lbu026.

Jaffe, A.B., Trajtenberg, M. (1998) International Knowledge Flows: Evidence From Patent Citations, Economics of Innovation and New Technology, 8:1-2, 105-136. DOI: 10.1080/10438599900000006.

Jaffe, A., Trajtenberg, M., Fogarty, M. (2000). Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors. The American Economic Review, 90(2), 215-218. Retrieved December 28, 2020, from http://www.jstor.org/stable/117223.

Laffi, M., Boschma, R. (2021). Does a local knowledge base in Industry 3.0 foster diversification in Industry 4.0 technologies? Evidence from European regions. Papers in Evolutionary Economic Geography. No.21.12. Utrecht University.

McCann, P. (2013) Modern Urban and Regional Economics. Oxford: Oxford University Press.

Maraut, S., Dernis, H., Webb, C., Spiezia, V., Guellec, D. (2008), "The OECD REGPAT database: a presentation", STI Working Paper 2008/2, OECD, Paris.

Mendonça, S. (2006). The revolution within: ICT and the shifiting knowledge base of the world's largest companies. Economics of Innovation and New Technology, 15:8, 777-799. DOI: 10.1080/10438590500510442.

Mendonça, S. (2009). Brave old world: Accounting for 'hightech' knowledge in 'low-tech' industries. Research Policy, Volume 38, Issue 3, pp. 470–482. https://doi.org/10.1016/j.respol.2008.10.018.

Montresor, S., Quatraro F. (2017). Regional Branching and Key Enabling Technologies: Evidence from European Patent Data, Economic Geography, 93:4, 367-396. DOI: 10.1080/00130095.2017.1326810.

Neffke, F., M. Henning, R. Boschma (2011), 'How do regions diversify over time? Industry relatedness and the development of new growth paths in regions,' Economic Geography, 87(3), 237–265. DOI:10.1111/ecge.2011.87.issue-3.

OECD, REGPAT database, January 2020.

OECD, Citations database, July 2020.

Perez, C. (2002). Technological Revolutions and Financial Capital: The Dynamics of Bubbles and Golden Ages. Cheltenham: Edward Elgar.

Rigby, D. (2015). Technological relatedness and knowledge space: Entry and exit of US cities from patent classes. Regional Studies. 49 (11): 1922–37. DOI:10.1080/00343404.2013.854878.

Rosenberg, N. (1979). Technological Interdependence in the American Economy. Technology and Culture, 20(1), 25-50. DOI:10.2307/3103110.

Schmoch, U. (2008). Updated in July 2019. "Concept of a technology classification for country comparisons". Final report to the World Intellectual Property Organisation (WIPO). WIPO. https://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/pdf/wipo_ipc_technology.pdf

Schumpeter, J. A. (1934). The theory of economic development. Cambridge, MA: Harvard University Press.

Tanner, A. N. (2016). The emergence of new technology-based industries: the case of fuel cells and its technological relatedness to regional knowledge bases. Journal of Economic Geography 16(3): 611-635. https://doi.org/10.1093/jeg/lbv011.

Weitzman, M. L. (1998). Recombinant Growth. Quarterly Journal of Economics. 113 (2): 331–60. https://doi.org/10.1162/003355398555595.

WIPO. (2019). Artificial Intelligence. Technology Trends 2019. Geneva, Switzerland: World Intellectual Property Organization.

WIPO. (2020). Artificial Intelligence. PATENTSCOPE Artificial Intelligence Index. Retrieved from https://www.wipo.int/tech_trends/en/artificial_intelligence/patentscope.html.

Xiao, J., Boschma, R., Andersson, M. (2018). Industrial Diversification in Europe: The Differentiated Role of Relatedness. Economic Geography. 94(5), 514-549. DOI: https://doi.org/10.1080/00130095.2018.1444989.

Appendix

Table A1. Correlation matrix of variables

Variables	(1)	(2)	(3)
AI_inventing (1)	1		
Ave_density (2)	0.3791	1	
Pop (3)	0.2019	0.3746	1