Papers in Evolutionary Economic Geography

21.12

Does a local knowledge base in Industry 3.0 foster diversification in Industry 4.0 technologies? Evidence from European regions

Matteo Laffi & Ron Boschma

Utrecht University Human Geography and Planning

Does a local knowledge base in Industry 3.0 foster diversification in Industry 4.0 technologies? Evidence from European regions

Matteo Laffi[1]

Ron Boschma[2,3]

[1]ABC Department, Politecnico di Milano

[2]Department of Human Geography and Planning, Utrecht University

[3]UiS Business School, University of Stavanger

7 March 2021

Abstract

The aim of the paper is to shed light on the role played by regional knowledge bases in Industry 3.0 in fostering new technologies in Industry 4.0 in European regions (NUTS3) over the period 1991-2015. We find that 4.0 technologies appear to be quite related to 3.0 technologies, with some heterogeneity among different technology fields. The paper investigates the geographical implications. We find that the probability of developing Industry 4.0 technologies is higher in regions that are specialised in Industry 3.0 technologies. However, other types of knowledge bases also sustain regional diversification in Industry 4.0 technologies.

JEL codes: B52, O33, R11

Keywords: Fourth Industrial Revolution, Industry 4.0, regional innovation, patents, knowledge space, relatedness, EU regions

1. Introduction

There is a long tradition of scholars pointing out that capitalist societies tend to go through a number of Industrial Revolutions (Perez and Soete 1988; Perez 2010). A well-established approach in the Economics of Innovation literature conceives the history of innovation as a temporal sequence of discrete jumps leading to new technological paradigms, followed by incremental changes along technological trajectories (Dosi 1982; Boschma 1999; Dosi and Nelson 2013). According to this view, the innovation process is characterized by historical moments in which the introduction of new and disruptive technologies (such as General-Purpose Technologies) paves the way for a complete re-organization of the economic system.

Scholars argue that we are now in the maturity phase of the ICT technological paradigm, also known as the technological paradigm 3.0, in which 3.0 technologies are now being improved, without significative discrete jumps. At the same time, scholars claim that the technological frontier is rapidly moving ahead, leading the economic system towards a fully-fledged Fourth Industrial Revolution (Brynjolfsson and Mcafee 2011, 2014; Schwab 2017). This so-called technological paradigm 4.0 is not characterized by a single and easily identifiable technology, but concerns a set of very different technologies (Ménière et al. 2017; Popkova et al. 2019). In particular, 4.0 technologies often combine advanced 3.0 technologies (both hardware and software) with technologies pertaining to different application domains. The disruptive novelty brought about by the 4.0 paradigm is that this process of recombination introduces radical change in fields that have not been extensively affected by the 3.0 paradigm in the past. Examples are the application of Internet of Things technologies in the agricultural sector or Artificial Intelligence tools in legal and business services (Liao et al. 2017; Lu 2017). Thus, on the one hand, we observe a certain degree of continuity between the 3.0 paradigm and 4.0 technologies, in which the 4.0 technological paradigm builds on the 3.0 paradigm through a process of continuous innovation. On the other hand, the broad recombinatory nature of the 4.0 paradigm introduces discontinuity in the innovation process: the application of advanced 3.0 technologies in new fields determines original and potentially disruptive 4.0 technical solutions that represent discrete jumps in the innovation process.

As happened with past industrial revolutions (Marshall 1987; Hall and Preston 1988; Boschma 1999), it is reasonable to expect that the transition from the 3.0 to the 4.0 paradigm will bring changes in the geography of innovation. This may depend on the distribution of typical material and immaterial inputs that are necessary for the creation of the particular technical knowledge on which the paradigm relies. In other words, there might be a connection between the technological features of each paradigm and its subsequent geography of innovation. Consequently, it is important to understand the mechanisms that influence the capacity of a region to play an active role in the creation of new technological knowledge. In the end, regional development depends not only on the adoption of technologies imported from outside the region (Boschma and Balland 2021a) but also on the possibility of producing new knowledge locally. Clearly, this is even more important when the unfolding of a new technological paradigm seems to be imminent. In this sense, regions face challenges and opportunities with huge impacts on their future development.

Despite the Fourth Industrial Revolution is attracting full attention, its geography and local determinants are still barely investigated. Among the few works on the topic (Gress and Kalafsky 2015; Strange and Zucchella 2017; Ciffolilli and Muscio 2018; Muscio and Ciffolilli 2020), the European Patent Office published a study on the geography of 4.0 innovation (Ménière et al. 2017), but it did not provide an analysis of its regional determinants. Balland and Boschma (2021b) identified European regions that display potential in the development of 4.0 technologies, showing the importance of relevant regional capabilities. They explored the connection between the local knowledge base and the ability of NUTS-2 regions in Europe to develop 4.0 technologies by applying the relatedness framework (Boschma 2017).

So, although the literature on the geography of 4.0 innovation is growing, there is still little evidence on their regional determinants. We analyze the ability of NUTS-3 regions in 32 European countries (EU-27, UK and the 4 EFTA countries) to develop new 4.0 technologies in the period 1991-2015. The present paper aims to move this literature a step forward with respect to at least three aspects. First, we shed new light on the connection between the two paradigms, exploring the degree of contuinity between the two from a relatedness framework. We do so by examing the degree of relatedness between 4.0 and 3.0 technologies, rather than providing a detailed technical overview of 4.0 technologies, which is beyond the scope of this paper. Our findings show that 4.0 technologies appear to be quite related to 3.0 technologies, with some heterogeneity among technological fields. Second, the paper explores the relationship between regional specialisation in 3.0 technologies and 4.0 knowledge creation in European NUTS-3 regions over the period 1991-2015. Our findings show that the probability of developing 4.0 technologies is higher in regions specialised in 3.0 technologies, especially for those 4.0 technologies that are closer to the 3.0 paradigm. Third, other types of technological specialisation are considered in order to test what kind of regional knowledge bases foster 4.0 knowledge creation. As recombinations lay at the heart of the 4.0 paradigm, we investigate whether local expertise and know-how in technologies coming from other fields may increase the ability of regions to diversify into 4.0 technologies.

The paper is organized as follows. Section 2 briefly discusses the relevant literature. Section 3 presents the data, while Section 4 and 5 explain the methodology adopted. Section 6 and 7 discuss the results and provide some robustness checks. Section 7 concludes.

2. Literature review

The long-term history of technology and innovation has been described in terms of a temporal sequence of discrete jumps and incremental changes. Dosi (1982, 1984; Dosi and Nelson 2013) developed an interpretative framework to describe this long-term pattern through the concepts of technological paradigms and technological trajectories. The advent of a new technological paradigm represents a discontinuity that sets into motion an incremental innovation process, which in turn take places along different technological trajectories. This framework has been further developed to include the phenomenon of Industrial Revolutions (Perez 2010). When some conditions are met, a new technological paradigm might have such disruptive consequences on the socio-economic system that the transition towards the new equilibrium represents a revolution.

Another feature of the innovation process that has drawn full attention is that it is subject to path dependence processes (Nelson and Winter 1982; Dosi et al. 1988; Arthur 1994). It is widely recognized that the existing knowledge base has a strong influence on paths of new knowledge creation and technological diversification. This influence is both direct and indirect. In the former case, the development of new technologies heavily relies and builds on existing technologies. In the latter, the availability of specific know-how, human capital, institutions, networks and peculiar resources that sustain the development of the existing knowledge base also shape future technological trajectories. From this point of view, the past and present technological structure conditions and shapes its future development (Castaldi and Dosi 2006).

The idea of path dependency has been applied in the geography literature, and is now considered to be one of the main pillars of Evolutionary Economic Geography (Boschma and Lambooy 1999; Martin and Sunley 2006; Henning et al. 2013). A more recent body of literature has empirically analyzed this idea of history matters using the relatedness concept to understand the diversification process in countries and regions (Boschma 2017). The main finding arising from these studies is that technological diversification in regions is more likely in those technological fields that are technologically “closer” to the ones in which the region already has relevant expertise (Kogler et al. 2013; Boschma et al. 2015; Rigby 2015). Unrelated diversification is rare and more likely to happen under particular conditions, such as unrelated variety (Castaldi et al. 2015), responsive institutions (Boschma and Capone 2015) and inflow of external actors (Neffke *et al.*, 2018).

The innovation considered in the present study is of a particular type, since it involves a change of technological paradigm. However, it is less clear whether the rise of the 4.0 paradigm represents a discontuity in the history of technology and innovation, or whether it shows signs of continuity in which processes of path dependency prevail. Some scholars emphasize its more radical and disruptive features, associating it with the Fourth Industrial Revolution (Brynjolfsson and Mcafee, 2011, 2014; Schwab 2017). Other scholars highlight the fact that 4.0 technologies combine technologies of the 3.0 paradigm with other technologies in particular application domains. This would imply that 3.0 technologies are being used and improved without significative discrete jumps, although, at the same time, they lead to radical change and disruptions in new fields of application (Liao et al. 2017; Lu 2017). This makes it crucial to determine whether 4.0 technologies and 3.0 technologies display some degree of cumulativeness, and which 4.0 technologies are closer to the ICT paradigm (continuity), and which ones are not (discontitnuity).

For a region, moving to 4.0 knowledge creation means to manage to exploit at its best all those elements that are necessary to reach the edge of the innovation frontier. Many regions in Europe show the ambition to do so (Santos et al. 2017; Reischauer 2018). However, there is little understanding which of them have the real potential to diversify into the technologies of the Fourth Industrial Revolution, as its geography and regional determinants are still underinvestigated (Gress and Kalafsky 2015; Strange and Zucchella 2017; Ciffolilli and Muscio 2018). The European Patent Office published a study with some evidence on the geography of 4.0 innovation (Ménière et al. 2017). However, this study did not provide an analysis of its regional determinants. Balland and Boschma (2021b) investigated the potential of NUTS-2 regions in Europe to develop 4.0 technologies. Using the analytical framework of relatedness (Boschma 2017), they found that regions were more likely to diversify into 4.0 technologies where they could draw on relevant local capabilities. However, what Balland and Boschma (2021b) did not investigate was the connection with 3.0 technologies, and to what extent there is discontinuity or continuity when shifting to a new paradigm from a geographical perspective.

The question has not yet been addressed whether a change of technological paradigm can be interpreted as a form of diversification in which regions leverage on their knowledge base and local inputs to develop new 4.0 technologies and "jump" into the new paradigm. The geographical implication of the possible relatedness between 3.0 and 4.0 technologies is that, in case of high levels of cumulativeness between the two paradigms, regions characterized by a well-developed 3.0 knowledge base might have an advantage in the production of 4.0 knowledge. Moreover, there is little understanding of what types of local knowledge bases are needed to develop what kinds of 4.0 technologies. Given the technological heterogeneity of the 4.0 paradigm, the local 3.0 knowledge base might have differentiated impacts on the development of different 4.0 technologies, depending on their level of cumulativeness with 3.0 technologies. What is more, the local know-how in a specific 4.0 technology might be a strategic asset for a region to diversify in other 4.0 technologies. This might be self-reinforcing: the higher the number of 4.0 technologies produced in a region, the easier it could be for that region to diversify in other 4.0 technologies. There is no evidence yet whether such a marginal effect exists. Furthermore, although 3.0 and 4.0 knowledge bases in regions are expected to be a strategic element in fostering 4.0 diversification, it is possible that other types of knowledge bases also play a role. As technologies coming from application fields are combined, a strong local knowledge base in these technologies could facilitate the diversification process in regions towards 4.0 technologies. Finally, analyses on 4.0 technologies have so far been done for NUTS-2 regions in Europe. We will look at the more detailed level of NUTS-3 regions instead, which enables us to analyze more precisely the importance of local capabilities.

3. Data on technologies

The study relies on patent data which is considered a good proxy for inventions and knowledge creation (Strumsky, Lobo and van der Leeuw, 2012; Strumsky and Lobo, 2015). The use of patent data allows identifying the regional knowledge base by looking at the technological codes that describe the patents produced in a region. Patent data are taken from the OECD Regpat database (Maraut *et al.*, 2008) and the inventions are regionalized according to the inventors' share. The geographical level of the analysis is the NUTS-3 European regions (EU27, the UK and the EFTA countries) and the overall time span considered is 1991-2015.

The first empirical challenge is to select among the CPC classification technological codes that represent 3.0 and 4.0 technologies. It is possible to argue that, given its phase of maturity, there is a broad consensus on the technological boundaries of the 3.0 paradigm. In order to select 3.0 technological patent codes, we adopted the classification of High-Tech IPC codes made by Eurostat (High-tech industry and knowledge-intensive services (htec), Annex 6) (Inaba and Squicciarini 2017). Two classes of ICT codes were selected, namely Computer and automated business equipment (ht_a) and Communication technologies (ht_f). These sets of IPC codes were mapped in the correspondent CPC codes using the official concordance tables.

Identifying a precise set of 4.0 technologies is not an easy task. There is a narrative of the Fourth Industrial Revolution based on anecdotes and examples of single technologies (e.g. artificial intelligence, 3D- printing, big data analytics, cloud computing, smart sensors). Only few studies provide a comprehensive description from a technological perspective, such as Chiarello et. al (2018) who exploited Wikipedia data to map clusters of Industry 4.0 technologies. A landmark study by the European Patent Office (Ménière et al. 2017) provides a sample of technological patent codes related to 4.0 technologies. This work is based on the expertise of technicians and patents examiners from the EPO. This paper makes use of the 4.0 technological codes provided by the EPO. Ménière et al. (2017) represents the most complete and reliable source available.

The EPOs' experts created a meaningful taxonomy of 4.0 inventions (patents) that allows assigning each technology to a specific class according to its characteristics. In particular, three macro technological classes were identified, each of them composed by some sub-categories, as reported in Table 1 in Annex A. The three classes are: Core technologies, Enabling technologies and Application domains. The class of Core technologies corresponds to the building blocks upon which 4.0 technologies are developed and include basic hardware technologies (sensors, processors, advanced memories), software technologies (adaptive databases, mobile operating systems, virtualisation) and connectivity systems (network protocols, adaptive wireless data systems). These are the advanced 3.0 technologies that can be recombined with other technologies in the context of the 4.0 technological paradigm. The second category, Enabling technologies, builds upon and complements the Core technologies paving the way for technological recombinations. Among Enabling technologies we find analytics systems, user interfaces (virtual reality), 3d technologies (printers and scanners), artificial intelligence (machine learning and neural networks), position determination systems (enhanced GPS), smart power supply technologies and intelligent safety systems. Finally, the Application domains refer to the final and recombinatory applications of 4.0 technologies in different parts of the economy, such as applications pertaining to individuals (wearables, health monitoring devices), application for the home environment (domotics), for moving vehicles, business enterprises (smart offices), manufacture (smart factories) and infrastructure.

The EPO study identified CPC codes belonging to at least one of the three macro classes. For the purpose of the study, the technological classes considered in the analysis are required to be mutually exclusive. Consequently, a new classification of 4.0 CPC codes is proposed, with 7 classes, as reported in Table 1. Each new class of 4.0 technologies corresponds to one of the three EPO classes, or to a mix of them. Classes 2, 3, 5 and 6 present a mix of the technological characteristics of macro-classes. If a technological code belongs both to the 3.0 definition and the 4.0 definition, the latter prevails and the CPC is considered as 4.0. As a robustness check, an alternative criterion was adopted in which only "pure" 3.0 and 4.0 CPC were considered, after having discarded those corresponding to both the definitions.

Table 1 – The new classification of 4.0 CPC codes based on EPO's classification

New classification of 4.0 technologies (CPC)		EPO's classification of 4.0 technologies (CPC)		
		Core tech.	Enabling tech.	Application dom.
1	Core technologies	1	0	0
2	Core and applied technologies	1	0	1
3	Core and enabling technologies	1	1	0
4	Enabling technologies	0	1	0
5	Enabling and applied technologies	0	1	1
6	Core, enabling and applied technologies	1	1	1
7	Applied technologies	0	0	1

Source: Authors' elaboration

4. Estimation of relatedness between technologies

The aim of the paper is to look at the local determinants of knowledge creation in 4.0 technologies. Following the relatedness framework, we expect the "closer" the knowledge base of a region is to 4.0 technologies, the more likely the region will develop 4.0 technologies. This requires a measure of technological distance between technologies. Following literature (Boschma, Balland and Kogler 2015; Rigby 2015), we computed a so-called "knowledge space", namely a kxk matrix - with k equals to the number of technologies – in which each element of the matrix is a standardized measure of the frequency with which the two technologies considered (i.e. two CPC codes) co-occur in a single patent in the sample of patents considered. The knowledge space was computed for 4 non-overlapping periods: 1991-1996;1997-2002;2003-2008;2009-2015. The calculations were made by exploiting some Python functions based on the EconGeo R package (Balland 2017).

Figure 1 presents the relatedness between the seven classes of 4.0 technologies and the two classes of 3.0 technologies for the last period: the lighter the colour, the higher the relatedness is between two technologies. 4.0 Core technologies (in particular class 1, 2 and 3) are on average the closest to the 3.0 class of Communication technologies (ht_f). This supports the interpretation of 4.0 Core technologies as very advanced 3.0 technologies: 4.0 Core technologies and 3.0 Communication technologies are often combined in the same invention. Interestingly, the 4.0 applied technologies (class no. 7) displays also high levels of relatedness with the 3.0 class of Computer and automated business equipment (ht_a) which suggests that this kind of technologies are somehow complementary for the realization of applied 4.0 solutions.

Figure 1 – Relatedness between 4.0 technologies and 3.0 technologies

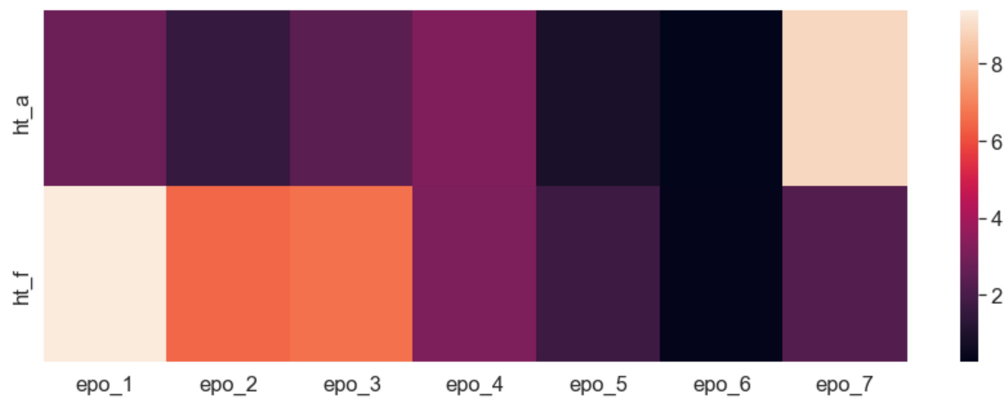
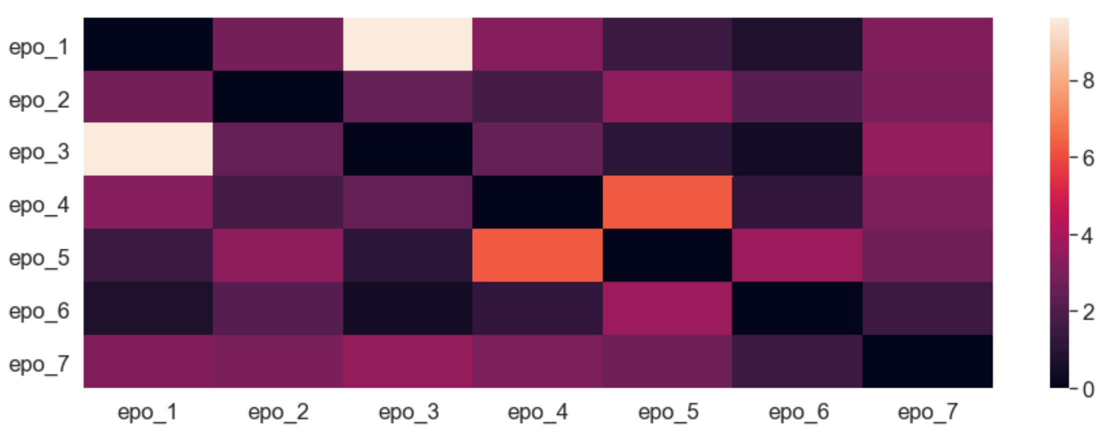


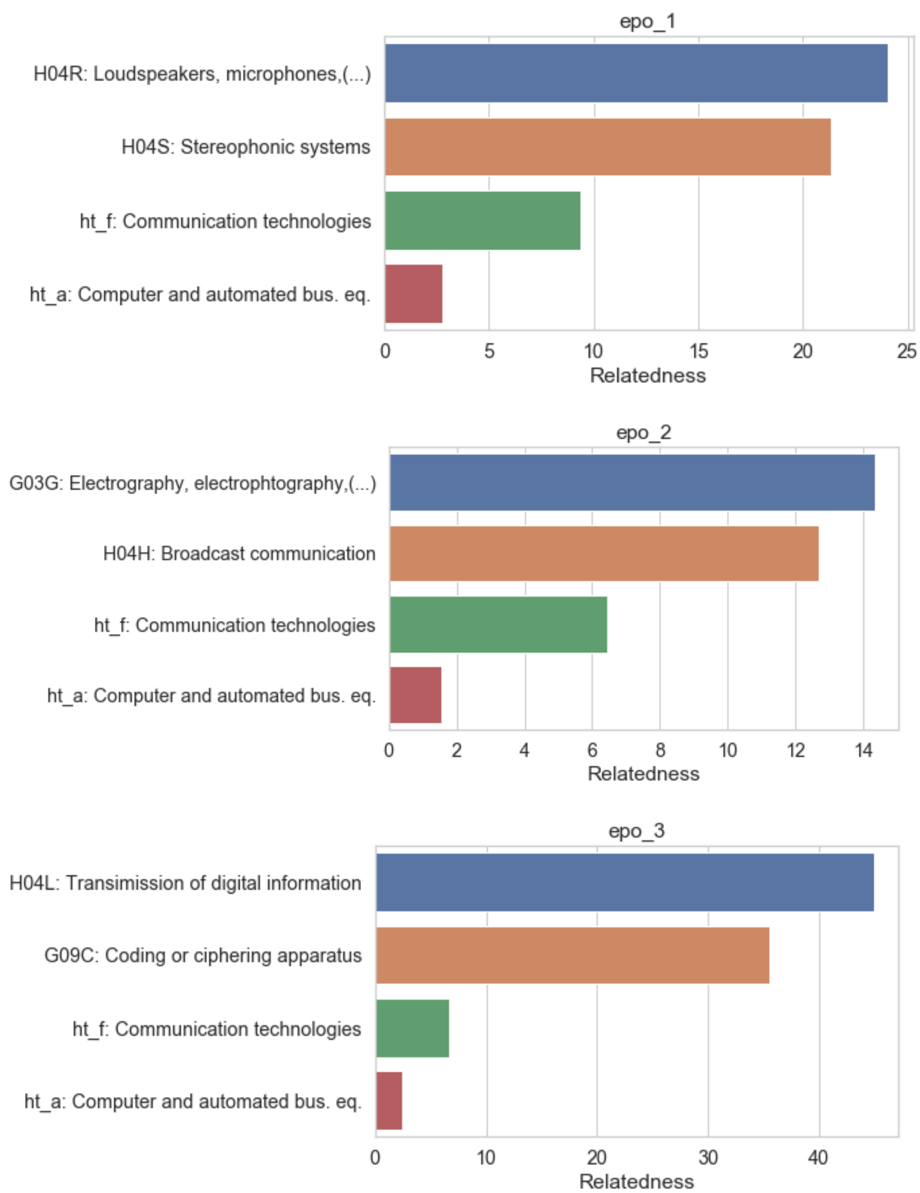
Figure 2 shows the relatedness between 4.0 technologies. Some of them are highly related, meaning that they frequently co-occur in the same patent, such as classes 1 and 3, and 4 and 5.

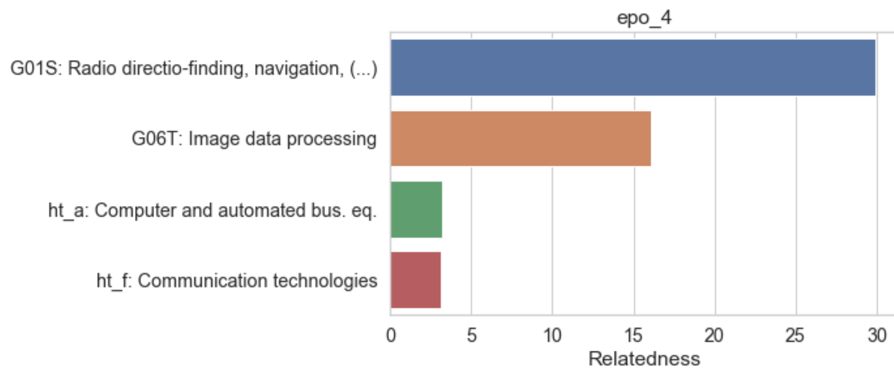
Figure 2 – Relatedness between 4.0 technologies

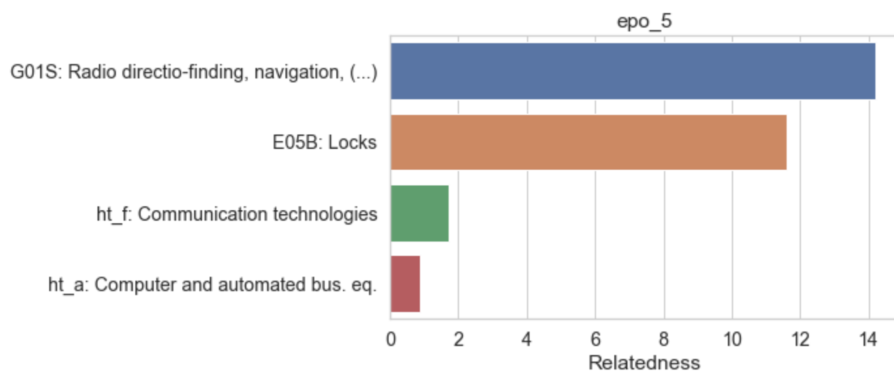


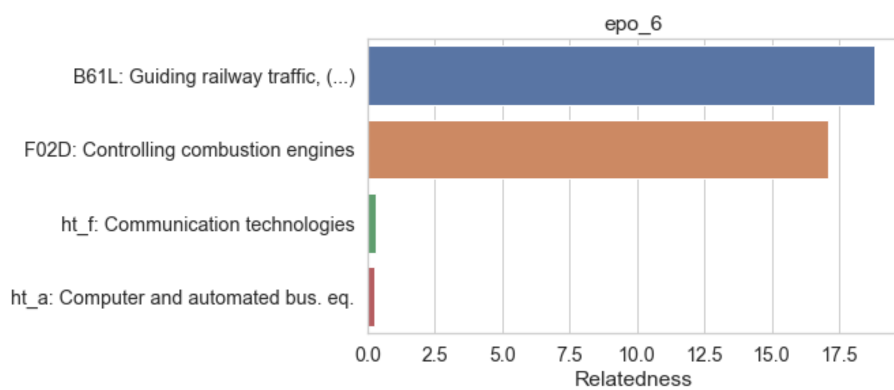
Do 4.0 technologies display higher levels of relatedness with technologies other than 3.0 ones? Figure 3 compares, for each of the seven classes of 4.0 technologies, the relatedness values of the top 2-related technologies to the relatedness values of the top 2-related 3.0 classes. What Figure 3 shows is that for none of the seven 4.0 classes, 3.0 technologies are among the 2 top-related technologies: other technologies are characterized by significantly higher levels of relatedness.

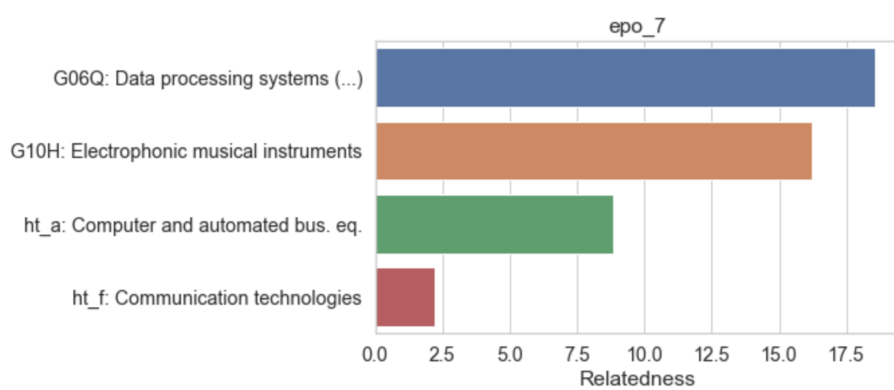
Figure 3 – Top 4.0 related technologies




epo_4
G01S: Radio directio-finding, navigation, (...)
G06T: Image data processing
ht_a: Computer and automated bus. eq.
ht_f: Communication technologies
0
5
10
15
20
25
30
Relatedness


epo_5
G01S: Radio directio-finding, navigation, (...)
E05B: Locks
ht_f: Communication technologies
ht_a: Computer and automated bus. eq.
0
2
4
6
8
10
12
14
Relatedness


epo_6
B61L: Guiding railway traffic, (...)
F02D: Controlling combustion engines
ht_f: Communication technologies
ht_a: Computer and automated bus. eq.
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Relatedness


epo_7
G06Q: Data processing systems (...)
G10H: Electrophonic musical instruments
ht_a: Computer and automated bus. eq.
ht_f: Communication technologies
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Relatedness

5. Modelling the entry of 4.0 technologies in European regions

To assess the importance of local determinants for the entry of new 4.0 technologies in a region, we estimate linear probability models based on a panel data structure with fixed effects for 4.0 technologies and time periods. Following studies on regional diversification, all observations include 4.0 technologies in which a region is not specialized (measured as a relative technological advantage (RTA) <1). The dependent variable is a dummy variable, $entry_{r,i,t}$, that takes value 1 when region *r* develops an RTA higher than 1 in a 4.0 technology i at time t, and 0 otherwise. Different specifications of the entry models aim at exploring different aspects that influence the probability of a region to develop new 4.0 technologies.

Model A – Baseline specification

The baseline specification of the model aims at verifying a standard result of the regional diversification literature in the context of 4.0 knowledge creation. More specifically, we test whether the probability to develop an RTA in a specific 4.0 technology is higher for those regions that are characterized by a local knowledge base close to that 4.0 technology. To do so, we calculated the relatedness density around the seven classes of 4.0 technologies (Balland et al. 2019) for each region. The relatedness density around a specific 4.0 technological class *i* in region *r* at time *t* is defined as the sum of technological relatedness $\varphi_{i,j,t}$ of technology i to all other technologies *j* (4.0 or not 4.0) in which the region has a RTA, divided by the sum of technological relatedness of technology *i* to all other technologies *j* in the reference region (i.e. the EU28 + EFTA countries) at time *t*:

$$Relatedness\_density_{r,i,t} = \frac{\sum_{j\in r, j\neq i} RTA_{r,j,t}\varphi_{i,j,t}}{\sum_{j\neq i}\varphi_{i,j,t}} * 100$$

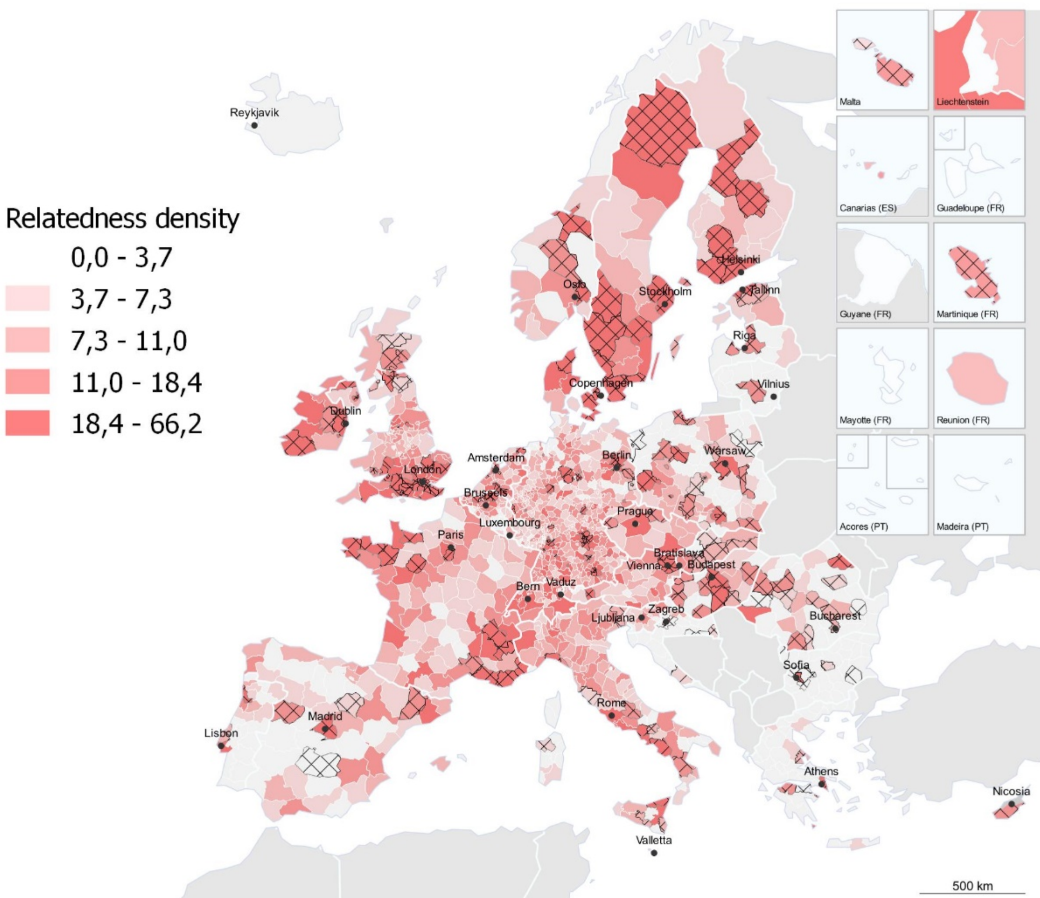
A region *r* has a $RTA_{r,j,t}$ in technology *j* at time *t* when the share of patents in technology *j* at time *t* in the region is greater than the same share in the reference area.

$$\begin{array}{ll} RTA_{r,j,t} = 1 \quad if & \dfrac{patents^t_{r,j}/\sum_j patents^t_{r,j}}{\sum_r patents^t_{r,j}/\sum_r\sum_j patents^t_{r,j}} > 1 \\ RTA_{r,j,t} = 0 & \qquad otherwise \end{array}$$

Thus, the higher is the value of relatedness density with respect to a certain 4.0 technology, the closer that technology is to the regional knowledge base.

To get a first indication, we compare in Figure 4 regions with an RTA in 4.0 technologies (indicated with a black grid pattern) and regions with high values of relatedness density (the darker the colour the higher the value) with respect to Core 4.0 technologies in period 4. What can be observed is that regions specialised in 4.0 technologies are also characterized by a high relatedness density. Moreover, also surrounding regions often tend to display high relatedness density values.

Figure 4 – Regional specialization in Core 4.0 technologies (black grid) and relatedness density (field colour) around Core 4.0 technologies (period 4)



Another way to look at the possible connection between relatedness density and the development of a specialization in 4.0 technologies is to classify regions in four category. Table 2 classifies the regions according to their level of relatedness density (values lower or higher than the median value) and their specialization in Core technologies (RTA lower or higher than 1) in period 4. 204 out of the 242 regions (84%) that present a 4.0 specialization in Core technologies display also higher than median values of relatedness density to 4.0 Core technologies. On the contrary, 656 out of the 1,146 regions (57%) that are not specialized in the production of Core technologies are characterized by lower than median values of relatedness density to Core technologies. It is interesting to note that many capital regions (for example Wien, Berlin, Madrid, Paris, Budapest) are characterized by both a specialisation and high levels of relatedness with respect to Core technologies. Instead, among the regions with a high relatedness density but without a specialization we can find advanced regions like Utrecht, Milan, Bergamo and Birmingham.

Table 2 – Relatedness density and specialization in Core technologies: a classification of European NUTS3 regions (period 4)

	Specialisation (RTA>1)	**No specialization (RTA<0)**
High relatedness density (> median value)	204	490
Low relatedness density (< median value)	38	656

In order to provide more systematic evidence, we conducted econometric analysis. Our entry model includes relatedness density as the main independent variable of interest. In more details:

$$(1) \qquad entry_{r,i,t} = \alpha_{r,i,t} + \beta_1 rel_{r,i,t-1} + \gamma_1 pop\_dens_{r,t} + \gamma_2 gdp_{r,t} + \theta_i + \mu_t + \varepsilon_{r,i,t}$$

where $rel_{r,i,t-1}$ represents the value of the relatedness density in region *r* for 4.0 technology *i* at time *t-1*. The coefficient β_1 is expected to be positive and significant, indicating a positive effect of relatedness to 4.0 technology *i* in the precedent period on the probability of developing an RTA in that technology in the following period. Two controls variables are included, namely regional population density ($pop\_dens_{r,t}$) and the regional gdp ($gdp_{r,t}$, calculated in pps). Both variables are considered at the beginning of the period and derived from Cambridge Econometrics[1]. Finally, θ_i and μ_t represent fixed effects at the technological and temporal level.

Model B – Effect of 3.0 technologies

To test the impact of regional specialization in 3.0 technologies, we added the two variables Computer and automated business equipment ($ht\_a_{r,t-1}$) and Communication technologies ($ht\_f_{r,t-1}$) that represent 3.0 technologies, and which take value of 1 when the region has an RTA>1 in the respective 3.0 technology at time *t-1*, and 0 otherwise. Appendix Annex C1 presents maps of European NUTS3-regions with respect to their scores on the two 3.0 technologies. A positive and significant value of coefficients β_2 and β_3 would suggest that, *ceteris paribus*, regions specialised in 3.0 technologies are more likely to develop 4.0 technologies. This takes the following form:

$$(2) \quad entry_{r,i,t} = \alpha_{r,i,t} + \beta_1 rel_{r,i,t-1} + \beta_2 ht\_a_{r,t-1} + \beta_3 ht\_f_{r,t-1} + \gamma cntr_{r,t} + \theta_i + \mu_t + \varepsilon_{r,i,t}$$

Model specification 3 adds two interaction terms that explore whether the effect of a specialization in 3.0 technologies on the probability of developing 4.0 technologies is greater for those 4.0 technologies that are technologically closer to the 3.0 paradigm:

$$(3) \quad entry_{r,i,t} = \alpha_{r,i,t} + \beta_1 rel_{r,i,t-1} + \beta_2 ht\_a_{r,t-1} + \beta_3 ht\_f_{r,t-1} + \beta_4 ht\_a\_int_{\ r,t-1} + \beta_5 ht\_f\_int_{r,t-1} + \gamma cntr_{r,t} + \theta_i + \mu_t + \varepsilon_{r,i,t}$$

where:

$$ht\_a\_int_{\ r,t-1} = epo123 * ht_{a_{r,t-1}}$$

$$ht\_f\_int_{\ r,t-1} = epo123 * ht\_f_{r,t-1}$$

and $epo123$ is a dummy variable taking value 1 when the observation concerns Core, Core and applied, and Core and enabling 4.0 technologies (classes 1,2,3).

[1] Source: https://www.camecon.com/european-regional-data/

Model C – Effect of other 4.0 technologies

We also estimate the possible effect of existing regional specialisations in 4.0 technologies on the probability of developing a new specialisation in another 4.0 technology. The analysis is based on model A and includes two dummy variables $few\_other40_{r,t-1}$ and $many\_other40_{r,t-1}$ which take value 1 if the region had a specialisation (RTA>1) in, respectively, one or two 4.0 technologies, and in more than two 4.0 technologies, different from i at time *t-1*. In this way, we verify whether the marginal effect of a 4.0 specialisation on the probability of developing a new specialisation in another 4.0 technology is increasing with the number of present 4.0 specialisations. In this case, we would have $\beta_7 > \beta_6$. Appendix Annex C2 provides a map of all European NUTS3 regions scoring on the number of 4.0 technologies in which they are specialised.

$$(4) \quad entry_{r,i,t} = \alpha_{r,i,t} + \beta_1 rel_{r,i,t-1} + \beta_2 ht\_a_{r,t-1} + \beta_3 ht\_f_{r,t-1} + \beta_6 few\_other40_{r,t-1} + \beta_7 many\_other40_{r,t-1} + \gamma cntr_{r,t} + \theta_i + \mu_t + \varepsilon_{r,i,t}$$

Model D – Effect of 'top-related' technologies

Model D includes the variable $toprel_{r,t-1}$, which captures the specialization of region *r* in the so-called top 4.0 related technologies. More precisely, $toprel_{r,t-1}$ takes value 1 when the region is specialised in at least one of the two top-related technologies with respect to 4.0 technology *i*. The relatedness density variable is excluded from this specification because it is highly correlated with this variable $toprel_{r,t-1}$, leading to an endogeneity problem.

$$(5) \quad entry_{r,i,t} = \alpha_{r,i,t} + \beta_2 ht\_a_{r,t-1} + \beta_3 ht\_f_{r,t-1} + \beta_4 toprel_{r,t-1} + \gamma cntr_{r,t} + \theta_i + \mu_t + \varepsilon_{r,i,t}$$

6. Results

Table 3 reports the results of the estimations of all models. First, the coefficient of the relatedness density variable is positive and highly significant in all specifications. This result confirms that the relatedness framework also holds in the context of 4.0 knowledge creation (Balland and Boschma 2021b): the probability of developing a specialization in a 4.0 technology is higher in those NUTS3 regions characterized by a knowledge base technologically close to that 4.0 technology.

Table 3 – Estimation results

Dependent variable: Entry(r,i,t)

	A	B1	B2	C	D
Reldens	0.00530***	0.00466***	0.00452***	0.00380***	
	(0.000427)	(0.000437)	(0.000439)	(0.000406)	
ht_a		-0.00740	-0.00812	-0.00785	0.00986
		(0.00814)	(0.0118)	(0.00793)	(0.00808)
ht_f		0.0743***	0.0536***	0.0590***	0.0918***
		(0.00933)	(0.0120)	(0.00840)	(0.00931)
ht_a_int			0.00310		
			(0.0157)		
ht_f_int			0.0534**		
			(0.0189)		
f_other40				0.0318***	
				(0.00587)	
m_other40				0.0667***	
				(0.00937)	
Toprel					0.0555***
					(0.0105)
Gdp	-0.00000204***	-0.00000194***	-0.00000187***	-0.00000203***	-0.000000554
	(0.000000391)	(0.000000391)	(0.000000392)	(0.000000371)	(0.000000367)
pop dens	0.00000194*	0.00000186*	0.00000187*	0.00000171*	0.00000190*
	(0.000000901)	(0.000000893)	(0.000000894)	(0.000000762)	(0.000000891)
N	17,938	17,938	17,938	17,938	17,938
R2	0.027	0.032	0.032	0.035	0.026

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Models B1 and B2 analyse the role of 3.0 regional specializations in fostering 4.0 technologies. The results of model B1 highlight a positive and significant effect on the entry probability only for Communication technologies (ht_f) but not for Automated business equipment (ht_a). It is worth noting that the magnitude of the coefficient is particularly relevant (0.0743), more than ten times bigger than the effect of relatedness on the entry probability (0.00466). Furthermore, model B2 tells us that this effect is even greater when the probability of developing specific types of 4.0 technologies is measured. Indeed, when considering Core 4.0 technologies, the impact of specialization in Communication technologies is much higher than the overall effect (0.0534+0.0536=0.107). These results confirm our expectations on the possible implications of technological cumulativeness between 3.0 and 4.0 technologies for the geography of 4.0 innovation. A local knowledge base in Communication technologies allows the region to leverage on the cumulative dimension of the 4.0 technological paradigm and, consequently, to produce more easily those kinds of 4.0 technologies that are more related to the 3.0 paradigm.

Model C demonstrates that having a specialization in some 4.0 technologies increases the probability of developing additional specializations in other classes of 4.0 technologies. Moreover, the higher the number of existing 4.0 specialisations, the easier it is for the region to diversify in a new 4.0 technology. Both coefficients of variables $few\_other40_{r,t-1}$ and $many\_other40_{r,t-1}$ are positive and significant, with the latter being higher than the former.

Model D shows how also a regional specialisation in those technologies that display the highest levels of 4.0 relatedness foster regional 4.0 innovation. Also in this case, the coefficient is positive, significant and high in magnitude (0.0555). This suggests that regional specialisation in 3.0 technologies is not the only driver of the development of 4.0 technologies.

Looking at the control variables, Table 3 shows that the coefficient of GDP is negative and usually significant, although the magnitude is barely negligible. Similarly, the positive, significant coefficient of population density does not provide relevant additional information.

7. Robustness checks

We performed a number of robustness checks. First, we included a time-invariant version of the regional 3.0 specialization variables. The rationale behind this choice is that, with the unfolding of the 4.0 technological paradigm, also those technologies classified as 3.0 might evolve in a similar direction, becoming more related to at least some of the 4.0 technologies. For this reason, we considered only the 3.0 specialization in the first period, when the 4.0 phenomenon was not present yet. The results in Table 4 are in line with the previous ones, with the exception of the interaction term for Communication technologies (ht_f_int_p1) which is not more significant. This evidence somehow supports this intuition, given that the link between a specialization in 3.0 technologies and 4.0 innovation in Core 4.0 technologies appears to be weaker when only the "old" specialization is considered, disregarding the subsequent evolution of the regional knowledge base.

Table 4 – Estimation results, 3.0 specialization variables calculated in period 1 (time-invariant)

Dependent variable: Entry(r,i,t); Specialisation in period 1

	B1	B2	C	D
relatedness	0.00482***	0.00477***	0.00380***	
	(0.000438)	(0.000439)	(0.000407)	
ht_a_p1	-0.00232	-0.00542	-0.00307	0.0183
	(0.00972)	(0.0139)	(0.00924)	(0.00964)
ht_f_p1	0.0581***	0.0471***	0.0460***	0.0740***
	(0.00919)	(0.0124)	(0.00834)	(0.00920)
ht_a_int_p1		0.00764		
		(0.0188)		
ht_f_int_p1		0.0265		
		(0.0183)		
f_other40			0.0335***	
			(0.00586)	
m_other40			0.0750***	
			(0.00918)	
Toprel				0.0554***
				(0.0105)
Gdp	-0.00000207***	-0.00000204***	-0.00000214***	-0.000000627
	(0.000000391)	(0.000000392)	(0.000000371)	(0.000000368)
pop dens	0.00000193*	0.00000194*	0.00000175*	0.00000197*
	(0.000000897)	(0.000000898)	(0.000000762)	(0.000000895)
N	17938	17938	17938	17938
R2	0.030	0.030	0.034	0.023

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Second, we considered an alternative specification of the dependent variable by adding a further condition on the regional development of an RTA in 4.0 technologies. The entry variable, in this case, takes value 1 only when (a) the location quotient becomes greater than 1 and (b) the absolute increase of the share of 4.0 technologies is larger than the minimum threshold of 0.3. Tables 1 and Table 2 in the Appendix, Annex D show the results and confirm our previous findings. Third, the sample of the regions considered in the analysis was filtered in order to exclude those cases with less than 10 patents per period. Table 3 and Table 4 in the Appendix, Annex D present the results and show the validity of our previous findings. Fourth, we changed the measurement of the 3.0 and 4.0 technologies (CPC). We excluded technologies that fitted both the 3.0 and the 4.0 definitions from our analysis (which applied to 1,795 CPC's, out of a total of 10,490 CPC's identified as 4.0), in order to exclude potential bias due to the partial overlapping of the 3.0 and 4.0 definitions. The results of the estimations are reported in Tables 5 and Table 6 in the Appendix, Annex D. They confirm earlier findings. Finally, we tested another specification of the dependent variable, namely an entry variable which takes value 1 when the new 4.0 technology has a RTA>2, instead of RTA>1. The findings are reported in Tables 7 and Table 8 in the Appendix, Annex D. They show the same results.

8. Conclusion and discussion

The paper aimed to shed light on the fundamental question what kinds of local knowledge bases have enabled the development of new Industry 4.0 technologies in Europe in the last 3 decades. First, we explored the extent to which local knowledge bases in 3.0 technologies laid the foundations of the development of 4.0 technologies in European regions at a very detailed level (the NUTS3 level). Second, we examined which other types of knowledge bases may have contributed to the development of new 4.0 technologies, applying recent insights from the empirical literature on regional diversification (Boschma 2017). Both questions are part of a much broader debate about possible links between Industry 3.0 and Industry 4.0 technologies (Brynjolfsson and Mcafee 2011; Schwab 2017). It centres around the key question whether Industry 4.0 stands for a major technological transformation that reflects a radical departure from existing technologies in general, and 3.0 technologies more in particular. Looking at this debate from a geographical lense, and adopting a relatedness framework, may provide new inputs to this.

First of all, we found that the knowledge space involving 4.0 technologies shows that the relationship between 4.0 technologies and 3.0 technologies is quite heterogeneous, with some 4.0 technologies being technologically closer to the previous 3.0 technological paradigm than others. Thus, there exists a certain degree of cumulativeness between the two paradigms, at least with respect to some 4.0 technologies. This cumulative dimension has some important implications for the resulting geography of Industry 4.0 innovation in Europe. In fact, the analysis showed that the probability of developing 4.0 technologies is higher in those regions that are specialised in the production of 3.0 technologies. This link is even stronger for the development of those 4.0 technologies that are closer to the 3.0 technological paradigm. At the same time, it is shown that the development of 4.0 technologies is a self-reinforcing process: when a region develops a specialisation in some 4.0 technologies, the probability of diversification in new 4.0 technologies increases, and this increase is larger when the number of 4.0 specialisations is higher.

The present study provides contributions but also opens the way for future research. First, it is important to take into consideration other local conditions that could enhance 4.0 technologies, going beyond the knowledge base approach adopted here. Although present technological trends influence future regional development paths, other regional variables could play a role and push regions in certain technological trajectories. The presence of local universities and other knowledge infrastructure (Tanner 2014, 2016) but also university-industry linkages might be crucial here (D'Este et al. 2013; Reischauer 2018). Institutional settings may also be considered important (Boschma and Capone 2015). Second, it could be interesting to include in the analysis some degree of spatial heterogeneity: regions are not all alike and different territories might present different modes of 4.0 knowledge creation. This point needs to be addressed in future research. Third, we only looked at regional knowledge bases, but we did not account for knowledge links with other regions. Inter-regional knowledge linkages can provide access to complementary capabilities (Miguelez and Moreno 2018; Balland and Boschma 2021a) and might enhance the ability of regions to contribute to the development of new 4.0 technologies, a topic that is still relatively unexplored. Fourth, there is a need to look more closely at the fields of application of Industry 4.0 technologies in regions. What needs to be explored whether there is overlap between the geographies of 4.0 technology production and the geographies of 4.0 industrial application in

Europe (De Propris and Bailey 2020). Fifth, there is need to focus on the consequences of Industry 4.0 technologies for spatial inequalities in Europe. This may be due to the fact that 4.0 technologies are likely to be highly complex, and therefore may have a tendency to concentrate in space, creating new spatial inequalities (Balland and Rigby 2017; Balland et al. 2019). But also the role of big and powerful companies need to be investigated and assessed in this respect, as they dominate the development of some 4.0 technologies (Ménière et al. 2017). This is part of a much broader debate that revolves around the quasi-monopolistic power of giant companies that are heavily engaged in Industry 4.0, and the types of reponses against the negative downsides of Industry 4.0 that come from citizens and the political system in different countries and regions (Feldman et al. 2019). No doubt this will impact the extent to which, and what types of Industry 4.0 technologies will be produced and implemented. It remains to be seen what consequences that will have for the future geography of Industry 4.0.

References

Arthur, W.B. (1994) *Increasing Returns and Path Dependence in the Economy*, University of Michigan Press, Ann Arbor, MI.

Balland, P. A. (2017) 'Economic Geography in R: Introduction to the EconGeo Package', *SSRN Electronic Journal*. doi: 10.2139/ssrn.2962146.

Balland, P. and R. Boschma (2021a) Complementary inter-regional linkages and Smart Specialisation. An empirical study on European regions, *Regional Studies*, doi: 10.1080/00343404.2020.1861240

Balland, P. and Boschma, R. (2021b) 'Mapping the potentials of regions in Europe to contribute to new knowledge production in Industry 4.0 technologies', *Regional Studies*, forthcoming.

Balland, P-A and Rigby, D (2017), The geography of complex knowledge, *Economic Geography*, 93(1), 1–23.

Balland, P.A., R. Boschma, J. Crespo and D. Rigby (2019) Smart specialization policy in the EU: Relatedness, knowledge complexity and regional diversification, Regional Studies 53 (9), 1252-1268.

Boschma, R.A. (1999) The rise of clusters of innovative industries in Belgium during the industrial epoch. *Research Policy* 28, 853-71

Boschma, R. (2017) 'Relatedness as driver of regional diversification: a research agenda', *Regional Studies*, 51(3), pp. 351–364. doi: 10.1080/00343404.2016.1254767.

Boschma, R., Balland, P.-A. and Kogler, D. F. (2015) 'Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010', *Industrial and Corporate Change*, 24(1), pp. 223–250. doi: 10.1093/icc/dtu012.

Boschma, R.A. and Capone, G (2015) Institutions and diversification: Related versus unrelated diversification in a varieties of capitalism framework, *Research Policy*, 44, 1902–1914.

Boschma, R. A., and Lambooy, J. G. (1999) Evolutionary economics and economic geography. *Journal of Evolutionary Economics* 9:411–29.

Brynjolfsson, E. and Mcafee, A. (2011) *Race Against the Machines: How the Digital Revolution is Accelerating Innovation, Driving Productivity and Irreversibly Transforming Employment and the Economy*. Lexington, MA: Digital Frontier Press.

Brynjolfsson, E. and Mcafee, A. (2014) *The Second Machine Age: Work, Progress and Prosperity in a Time of Brilliant Technologies*. London, UK: W.W: Norton & Company.

Castaldi C. and Dosi G. (2006) The Grip of History and the Scope for Novelty: Some Results and Open Questions on Path Dependence in Economic Processes. In: Wimmer A., Kössler R. (eds) Understanding Change. Palgrave Macmillan, London, pp. 99-128, https://doi.org/10.1057/9780230524644_8

Castaldi, C, Frenken, K and Los, B (2015) Related variety, unrelated variety and technological breakthroughs. An analysis of US state-level patenting, *Regional Studies*, 49(5), 767–781.

Chiarello, F. *et al.* (2018) ‘Extracting and mapping industry 4.0 technologies using wikipedia’, *Computers in Industry*. Elsevier, 100(September 2017), pp. 244–257. doi: 10.1016/j.compind.2018.04.006.

Ciffolilli, A. and Muscio, A. (2018) ‘Industry 4.0: national and regional comparative advantages in key enabling technologies’, *European Planning Studies*, 26(12), pp. 2323–2343. doi: 10.1080/09654313.2018.1529145.

De Propris, L. and D. Bailey (Eds.) (2020) Industry 4.0 and Regional Transformations, London/New York: Routledge.

D'Este, P., F. Guy and S. Iammarino (2013) Shaping the formation of university–industry research collaborations: what type of proximity does really matter?, *Journal of Economic Geography* 13 (4), 537–558.

Dosi, G. (1982) 'Technological paradigms and technological trajectories', *Research Policy*, 11(3), pp. 147–162. doi: 10.1016/0048-7333(82)90016-6.

Dosi, G., Freeman C., Nelson R., Silverberg G., Soete, L. (1988) 'Technical Change and Economic Theory', Pinter Publisher, London and New York

Dosi, G. and Nelson, R. R. (2013) 'The Evolution of Technologies: An Assessment of the State-of-the-Art', *Eurasian Business Review*, (3), pp. 3–46. doi: 10.14208/BF03353816.

Feldman M., Guy F., Iammarino S. (2019), Regional Income Disparities, Monopoly & Finance, Working Paper. Birbeck, University of London, 43, October, pp. 1-30.

Gress, D. R. and Kalafsky, R. V. (2015) 'Geographies of production in 3D: Theoretical and research implications stemming from additive manufacturing', *Geoforum*, 60, pp. 43–52. doi: 10.1016/j.geoforum.2015.01.003.

Hall, P. and Preston, P. (1988) 'The Carrier Wave: New Information Technology and the Geography of Innovation, 1846-2003'. London: Unwin Hyman

Henning, M. E. Stam and R. Wenting (2013) Path dependence research in regional economic development: Cacophony or knowledge accumulation? *Regional Studies* 47 (8), pp. 1348-1362.

Hidalgo, C. A. *et al.* (2018) 'The Principle of Relatedness', in *Springer Proceedings in Complexity*, pp. 451–457. doi: 10.1007/978-3-319-96661-8_46.

Inaba, T. and Squicciarini, M. (2017) 'ICT: A new taxonomy based on the international patent classification', *OECD Science, Technology and Industry Working Papers*, 01. doi: https://dx.doi.org/10.1787/ab16c396-en.

Kogler, D.F., Rigby D.L., Tucker, I. (2013) 'Mapping knowledge space and technological relatedness in US cities'. *European Planning Studies* 21(9), 1374-1391.

Liao, Y. *et al.* (2017) 'Past, present and future of Industry 4.0 - a systematic literature review and research agenda proposal', *International Journal of Production Research*, 55(12), pp. 3609–3629. doi: 10.1080/00207543.2017.1308576.

Lu, Y. (2017) 'Industry 4.0: A survey on technologies, applications and open research issues', *Journal of Industrial Information Integration*, 6, pp. 1–10. doi: 10.1016/j.jii.2017.04.005.

Maraut, S. *et al.* (2008) 'The OECD REGPAT database: a presentation', *STI Working Paper*. doi: 10.1787/241437144144.

Marshall, M. (1987) *Long waves of regional development*, London: MacMillan.

Martin R. and Sunley P. (2006). 'Path dependence and regional economic evolution'. *Journal of Economic Geography* 6 (4), 395–437.

Ménière, Y., I. Rudyk and J. Valdes (2017), *Patents and the Fourth Industrial Revolution. The inventions behind digital transformation*, European Patent Office, Munich.

Miguelez, E. and R. Moreno (2018) Relatedness, external linkages and regional innovation in Europe, *Regional Studies* 52, 688–701.

Muscio, A. and A. Ciffolilli (2020) What drives the capacity to integrate Industry 4.0 technologies? Evidence from European R&D projects, *Economics of Innovation and New Technology* 29 (2), 169-183.

Neffke, F, Hartog, M, Boschma, RA and Henning, M (2018) Agents of structural change: The role of firms and entrepreneurs in regional diversification, *Economic Geography*, 94(1), pp. 23–48. doi: 10.1080/00130095.2017.1391691.

Neffke, F, Henning M and Boschma, RA (2011), How do regions diversify over time? Industry relatedness and the development of new growth paths in regions, *Economic Geography*, 87(3), 237–265.

Nelson, R. R. and Winter S. G. (1982). *An Evolutionary Theory of Economic Change*. (Cambridge, MA and London: The Belknap Press).

Perez, C. (2010) 'Technological revolutions and techno-economic paradigms', *Cambridge Journal of Economics*, 34(1), pp. 185–202. doi: 10.1093/cje/bep051.

Perez, C and Soete, L (1988) Catching up in technology: entry barriers and windows of opportunity, in G Dosi, C Freeman, R Nelson, G Silverberg & L Soete (eds), *Technical Change and Economic Theory*, Pinter Publishers, London, pp. 458-479.

Popkova, E. G., Y.V.Ragulina and A.V.Bogoviz (Eds.) (2019) *Industry 4.0: Industrial Revolution of the 21st Century*, Springer.

Reischauer, G. (2018), Industry 4.0 as policy-driven discourse to institutionalize innovation systems in manufacturing, *Technological Forecasting and Social Change* 132, pp. 26-33.

Rigby, D. L. (2015) 'Technological Relatedness and Knowledge Space: Entry and Exit of US Cities from Patent Classes', *Regional Studies*, 49(11), pp. 1922–1937. doi: 10.1080/00343404.2013.854878.

Santos, C., Mehrsai, A., Barros, A. C., Araújo, M., & Ares, E. (2017). Towards Industry 4.0: An overview of European strategic roadmaps. *Procedia Manufacturing*, *13*, 972–979.

Schwab, K. (2017) *The Fourth Industrial Revolution: what it means and how to respond*. Edited by P. B. Ltd. London, UK.

Strange, R. and Zucchella, A. (2017) 'Industry 4.0, global value chains and international business', *Multinational Business Review*, 25(3), pp. 174–184. doi: 10.1108/MBR-05-2017-0028.

Strumsky, D. and Lobo, J. (2015) 'Identifying the sources of technological novelty in the process of invention', *Research Policy*. Elsevier B.V., 44(8), pp. 1445–1461. doi: 10.1016/j.respol.2015.05.008.

Strumsky, D., Lobo, J. and van der Leeuw, S. (2012) 'Using patent technology codes to study technological change', *Economics of Innovation and New Technology*, 21(3), pp. 267–286. doi: 10.1080/10438599.2011.578709.

Tanner, A. N. (2014) Regional branching reconsidered: Emergence of the fuel cell industry in European regions. Economic Geography, 90(4), 403–427. doi:10.1111/ecge.12055

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. doi:10.1093/jeg/lbv011

Annex A

Table 1 – 4.0 technologies classes proposed by Valdes (2017)

Technological fields	Examples
Core technologies	
Hardware	Sensors, advanced memories, processors, adaptive displays
Software	Intelligent cloud storage and computing structures, adaptive databases, mobile operating systems, virtualization
Connectivity	Network protocols for massively connected devices, adaptive wireless data systems
Enabling technologies	
Analytics	Diagnostic systems for massive data
User interfaces	User interfaces, virtual reality, information display in eyewear
Three-dimensional support systems	Additive manufacturing, 3D printers and scanners for parts manufacture, automated 3D design and simulation
Artificial intelligence	Artificial intelligence, machine learning, neural networks
Position determination	Enhanced GPS, device to device relative and absolute positioning
Power supply	Situation-aware charging systems, shared power transmission objectives
Security	Adaptive security systems, intelligent safety systems
Application domains	
Personal	Personal health monitoring devices, smart wearables, entertainment devices
Home	Smart homes, alarm systems, intelligent lighting and heating, consumer robotics
Vehicles	Autonomous driving, vehicle fleet navigation devices
Enterprise	Intelligent retail and healthcare systems, autonomous office systems, smart offices, agriculture
Manufacture	Smart factories, intelligent robotics, energy saving
Infrastructure	Intelligent energy distribution networks, intelligent transport networks, intelligent lighting and heating systems

Source: Valdes (2017)

Annex B

Table 1 – List of the 2 top-4.0 related technologies

4.0 class	Top-rel. tech 1	Top-rel tech 2
Epo_1	H04R (LOUDSPEAKERS, MICROPHONES, GRAMOPHONE PICK-UPS OR LIKE ACOUSTIC ELECTROMECHANICAL TRANSDUCERS; DEAF-AID SETS; PUBLIC ADDRESS SYSTEMS)	H04S (STEREOPHONIC SYSTEMS)
Epo_2	G03G (ELECTROGRAPHY; ELECTROPHOTOGRAPHY; MAGNETOGRAPHY)	H04H (BROADCAST COMMUNICATION)
Epo_3	H04L (TRANSMISSION OF DIGITAL INFORMATION, e.g. TELEGRAPHIC COMMUNICATION)	G09C (CODING OR CIPHERING APPARATUS FOR CRYPTOGRAPHIC OR OTHER PURPOSES INVOLVING THE NEED FOR SECRECY)
Epo_4	G01S (RADIO DIRECTION-FINDING; RADIO NAVIGATION; DETERMINING DISTANCE OR VELOCITY BY USE OF RADIO WAVES; LOCATING OR PRESENCE-DETECTING BY USE OF THE REFLECTION OR RERADIATION OF RADIO WAVES; ANALOGOUS ARRANGEMENTS USING OTHER WAVES)	G06T (IMAGE DATA PROCESSING OR GENERATION, IN GENERAL)
Epo_5	G01S (Se above)	E05B (LOCKS; ACCESSORIES THEREFOR; HANDCUFFS)
Epo_6	B61L (GUIDING RAILWAY TRAFFIC; ENSURING THE SAFETY OF RAILWAY TRAFFIC)	F02D (CONTROLLING COMBUSTION ENGINES)
Epo_7	G06Q (DATA PROCESSING SYSTEMS OR METHODS, SPECIALLY ADAPTED FOR ADMINISTRATIVE, COMMERCIAL, FINANCIAL, MANAGERIAL, SUPERVISORY OR FORECASTING PURPOSES)	G10H (ELECTROPHONIC MUSICAL INSTRUMENTS)

Annex C1

Figure 1 – Regional specialisation in 3.0 technologies, Computer and Automated business equipment (ht_a, period 1)

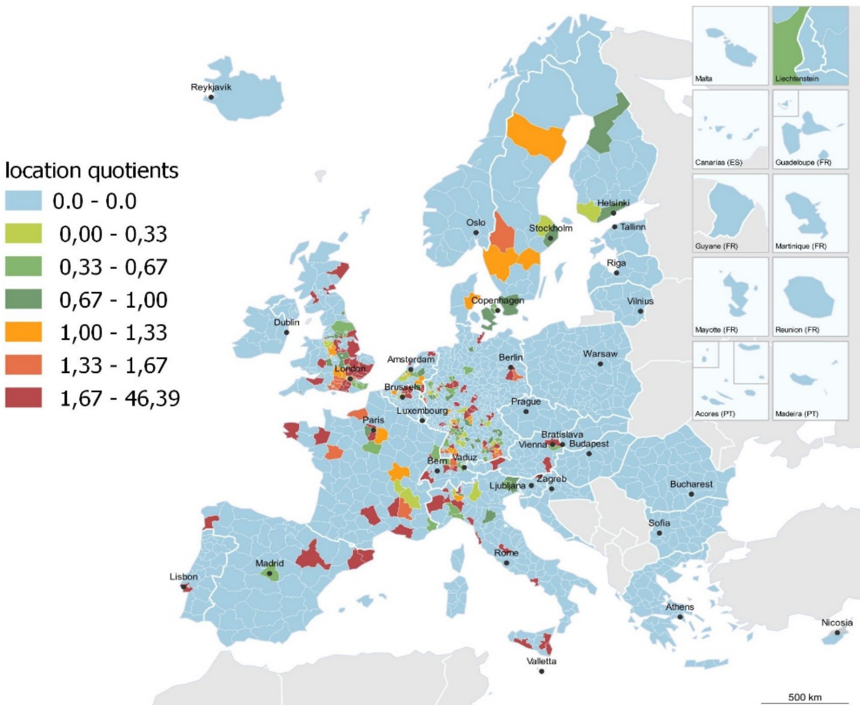
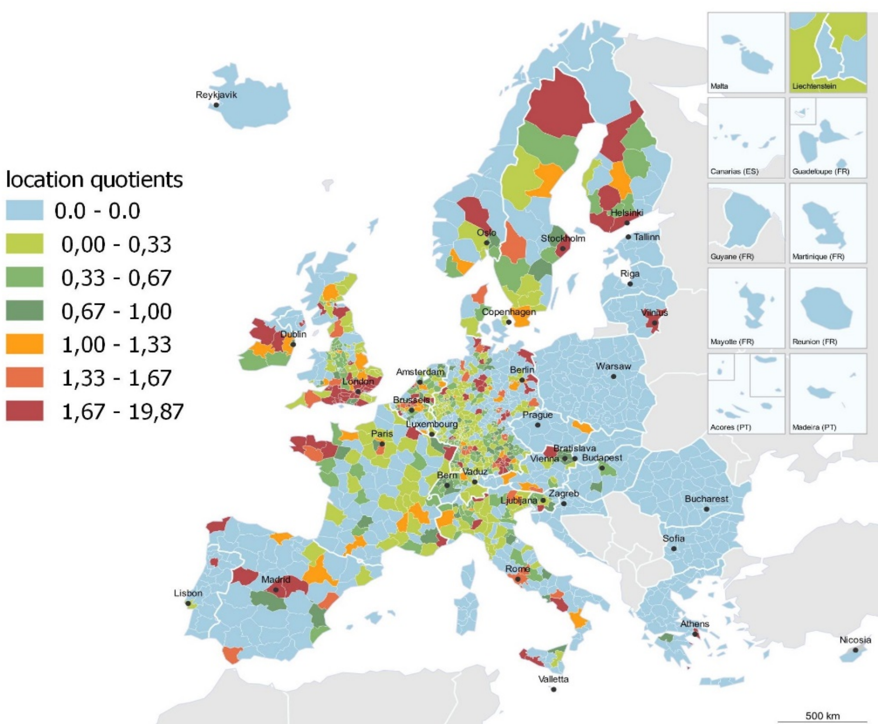
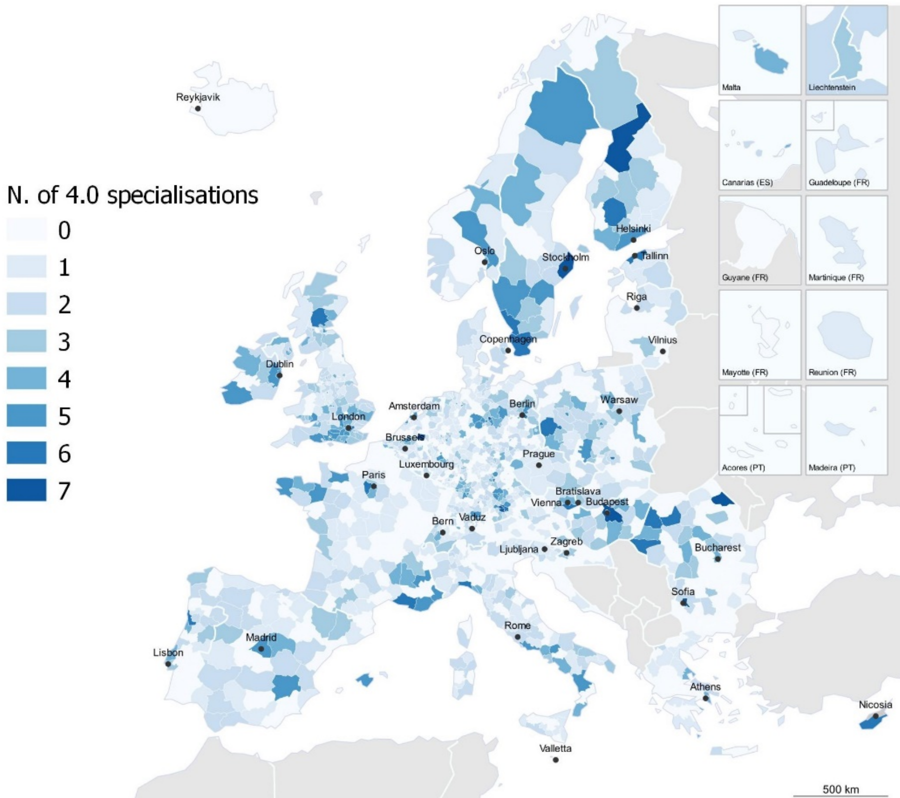


Figure 2 – Regional specialisation in 3.0 technologies, Communication technologies (ht_f, period 1)



Annex C2

Figure 1 – Regional specialisation in 4.0 technologies (number of classes in which a region has a specialisation, period 4)



Annex D

Table 1 – Robustness checks: at least 0.3 increase in the share of technological specialization for the dependent variable

Dependent variable: Entry(r,i,t)

	A	B1	B2	C	D
relatedness	0.00401***	0.00341***	0.00328***	0.00251***	
	(0.000410)	(0.000421)	(0.000423)	(0.000397)	
ht_a		-0.00545	-0.00686	-0.00604	0.00720
		(0.00794)	(0.0115)	(0.00776)	(0.00786)
ht_f		0.0681***	0.0497***	0.0530***	0.0809***
		(0.00911)	(0.0117)	(0.00821)	(0.00904)
ht_a_int			0.00446		
			(0.0153)		
ht_f_int			0.0476**		
			(0.0184)		
f_other40				0.0376***	
				(0.00572)	
m_other40				0.0675***	
				(0.00916)	
Toprel					0.0422***
					(0.0101)
Gdp	-0.00000182***	-0.00000174***	-0.00000167***	-0.00000185***	-0.000000714*
	(0.000000380)	(0.000000380)	(0.000000381)	(0.000000363)	(0.000000355)
pop dens	0.00000133	0.00000126	0.00000127	0.00000108	0.00000129
	(0.000000832)	(0.000000825)	(0.000000827)	(0.000000746)	(0.000000824)
N	18084	18084	18084	18084	18084
R2	0.020	0.024	0.024	0.027	0.020

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2 – Robustness checks: at least 0.3 increase in the share of technological specialization for the dependent variable

Dependent variable: Entry(r,i,t); Specialisation in period 1

	B1	B2	C	D
relatedness	0.00356***	0.00351***	0.00251***	
	(0.000422)	(0.000423)	(0.000398)	
ht_a_p1	0.000703	-0.00460	-0.000522	0.0160
	(0.00951)	(0.0136)	(0.00905)	(0.00939)
ht_f_p1	0.0513***	0.0392**	0.0396***	0.0631***
	(0.00895)	(0.0121)	(0.00814)	(0.00891)
ht_a_int_p1		0.0126		
		(0.0184)		
ht_f_int_p1		0.0293		
		(0.0178)		
f_other40			0.0391***	
			(0.00572)	
m_other40			0.0752***	
			(0.00898)	
toprel				0.0421***
				(0.0102)
pop dens	0.00000132	0.00000133	0.00000112	0.00000135
	(0.000000829)	(0.000000830)	(0.000000746)	(0.000000828)
gdp	-0.00000185***	-0.00000182***	-0.00000194***	-0.000000776*
	(0.000000380)	(0.000000381)	(0.000000362)	(0.000000356)
N	18084	18084	18084	18084
R2	0.022	0.022	0.026	0.018

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3 - Robustness check: at least 0.3 increase in the share of technological specialization for the dependent variable, only regions with more than 10 technologies

Dependent variable: Entry(r,i,t)

	A	B1	B2	C	D
relatedness	0.00340***	0.00281***	0.00265***	0.00207***	
	(0.000430)	(0.000440)	(0.000442)	(0.000413)	
ht_a		-0.00609	-0.00896	-0.00644	0.00342
		(0.00799)	(0.0116)	(0.00794)	(0.00793)
ht_f		0.0687***	0.0472***	0.0539***	0.0788***
		(0.00934)	(0.0120)	(0.00852)	(0.00926)
ht_a_int			0.00786		
			(0.0155)		
ht_f_int			0.0553**		
			(0.0189)		
f_other40				0.0312***	
				(0.00608)	
m_other40				0.0632***	
				(0.00943)	
toprel					0.0388***
					(0.0102)
gdp	-0.00000233***	-0.00000223***	-0.00000216***	-0.00000225***	-0.00000151***
	(0.000000404)	(0.000000404)	(0.000000405)	(0.000000382)	(0.000000387)
pop dens	0.00000119	0.00000112	0.00000112	0.000000974	0.00000112
	(0.000000837)	(0.000000830)	(0.000000833)	(0.000000763)	(0.000000828)
N	16776	16776	16776	16776	16776
R2	0.019	0.023	0.024	0.026	0.021

Standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

Table 4 - Robustness check: at least 0.3 increase in the share of technological specialization for the dependent variable, only regions with more than 10 technologies

Dependent variable: Entry(r,i,t); Specialisation in period 1

	B1	B2	C	D
relatedness	0.00298***	0.00292***	0.00208***	
	(0.000441)	(0.000442)	(0.000414)	
ht_a_p1	0.000740	-0.00530	-0.0000588	0.0127
	(0.00962)	(0.0137)	(0.00928)	(0.00949)
ht_f_p1	0.0496***	0.0350**	0.0383***	0.0587***
	(0.00904)	(0.0122)	(0.00836)	(0.00902)
ht_a_int_p1		0.0144		
		(0.0186)		
ht_f_int_p1		0.0353		
		(0.0180)		
f_other40			0.0329***	
			(0.00607)	
m_other40			0.0716***	
			(0.00924)	
toprel				0.0388***
				(0.0102)
pop dens	0.00000118	0.00000119	0.00000102	0.00000118
	(0.000000834)	(0.000000836)	(0.000000763)	(0.000000832)
gdp	-0.00000234***	-0.00000230***	-0.00000234***	-0.00000157***
	(0.000000404)	(0.000000405)	(0.000000382)	(0.000000388)
N	16776	16776	16776	16776
R2	0.021	0.021	0.025	0.019

Standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

Table 5 - Robustness check: at least 0.3 increase in the share of technological specialization for the dependent variable, only regions with more than 10 technologies, and stricter definition of 3.0 and 4.0 technologies (CPC)

Dependent variable: Entry(r,i,t)

	A	B1	B2	C	D
relatedness	0.00486***	0.00438***	0.00416***	0.00390***	
	(0.000423)	(0.000435)	(0.000437)	(0.000414)	
ht_a		0.00145	-0.00382	0.000601	0.0168*
		(0.00861)	(0.0121)	(0.00831)	(0.00850)
ht_f		0.0487***	0.0222	0.0443***	0.0604***
		(0.00916)	(0.0117)	(0.00838)	(0.00917)
ht_a_int			0.0140		
			(0.0167)		
ht_f_int			0.0675***		
			(0.0186)		
f_other40				0.0217**	
				(0.00663)	
m_other40				0.0359***	
				(0.00963)	
toprel					0.0602***
					(0.00960)
gdp	-0.00000265***	-0.00000262***	-0.00000251***	-0.00000264***	-0.00000148***
	(0.000000410)	(0.000000410)	(0.000000410)	(0.000000392)	(0.000000389)
pop dens	0.00000156	0.00000149	0.00000149	0.00000144	0.00000150
	(0.000000912)	(0.000000908)	(0.000000910)	(0.000000812)	(0.000000909)
N	16501	16501	16501	16501	16501
R2	0.037	0.039	0.040	0.040	0.034

Standard errors in parentheses
* p < 0.05, ** p < 0.01, *** p < 0.001

Table 6 - Robustness check: at least 0.3 increase in the share of technological specialization for the dependent variable, only regions with more than 10 technologies, and a stricter definition of 3.0 and 4.0 technologies (CPC)

Dependent variable: Entry(r,i,t); Specialisation in period 1

	B1	B2	C	D
relatedness	0.00460***	0.00452***	0.00406***	
	(0.000433)	(0.000434)	(0.000413)	
ht_a_p1	-0.00562	-0.0141	-0.00672	0.0127
	(0.0101)	(0.0141)	(0.00969)	(0.00998)
ht_f_p1	0.0378***	0.0211	0.0338***	0.0494***
	(0.00909)	(0.0122)	(0.00851)	(0.00913)
ht_a_int_p1		0.0200		
		(0.0196)		
ht_f_int_p1		0.0401*		
		(0.0182)		
f_other40			0.0227***	
			(0.00663)	
m_other40			0.0393***	
			(0.00959)	
toprel				0.0640***
				(0.00957)
pop dens	0.00000154	0.00000154	0.00000149	0.00000153
	(0.000000910)	(0.000000911)	(0.000000812)	(0.000000910)
gdp	-0.00000269***	-0.00000264***	-0.00000271***	-0.00000151***
	(0.000000410)	(0.000000411)	(0.000000392)	(0.000000391)
N	16501	16501	16501	16501
R2	0.038	0.038	0.039	0.033

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7 - Robustness check: RTA for the entry variable calculated with a threshold of 2

Dependent variable: Entry(r,i,t)

	A	B1	B2	C	D
relatedness	0.000874***	0.000532*	0.000408	0.000285	
	(0.000229)	(0.000235)	(0.000233)	(0.000213)	
ht_a		-0.0104*	-0.00863	-0.0106*	-0.00883*
		(0.00437)	(0.00610)	(0.00452)	(0.00426)
ht_f		0.0350***	0.0201**	0.0297***	0.0373***
		(0.00512)	(0.00656)	(0.00471)	(0.00511)
ht_a_int			-0.00276		
			(0.00833)		
ht_f_int			0.0379***		
			(0.0105)		
f_other40				0.00475	
				(0.00355)	
m_other40				0.0206***	
				(0.00535)	
toprel					0.0170**
					(0.00527)
gdp	-0.000000994***	-0.000000971***	-0.000000919***	-0.00000100***	-0.000000867***
	(0.000000220)	(0.000000220)	(0.000000220)	(0.000000209)	(0.000000211)
pop dens	0.000000709	0.000000623	0.000000628	0.000000588	0.000000620
	(0.000000500)	(0.000000495)	(0.000000497)	(0.000000464)	(0.000000498)
N	21056	21056	21056	21056	21056
R2	0.004	0.007	0.008	0.008	0.008

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8 - Robustness check: RTA for the entry variable calculated with a threshold of 2

Dependent variable: Entry(r,i,t); Specialisation in period 1

	B1	B2	C	D
relatedness	0.000554*	0.000521*	0.000242	
	(0.000234)	(0.000235)	(0.000214)	
ht_a_p1	-0.00498	-0.00943	-0.00495	-0.00319
	(0.00514)	(0.00699)	(0.00521)	(0.00509)
ht_f_p1	0.0323***	0.0263***	0.0277***	0.0344***
	(0.00520)	(0.00692)	(0.00471)	(0.00516)
ht_a_int_p1		0.0108		
		(0.0101)		
ht_f_int_p1		0.0146		
		(0.0104)		
few_other40			0.00548	
			(0.00355)	
many_other40			0.0236***	
			(0.00525)	
toprel				0.0170**
				(0.00529)
pop dens	0.000000688	0.000000691	0.000000640	0.000000684
	(0.000000498)	(0.000000499)	(0.000000464)	(0.000000500)
gdp	-0.00000103***	-0.00000101***	-0.00000106***	-0.000000915***
	(0.000000220)	(0.000000221)	(0.000000209)	(0.000000212)
N	21056	21056	21056	21056
R2	0.007	0.007	0.008	0.007

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$