



Scientific knowledge production in European regions: patterns of growth, diversity and complexity

Gaston Heimeriks ^{a,b}, Deyu Li^a, Wout Lamers^b, Ingeborg Meijer^b and Alfredo Yegros^b

^aInnovation Studies, Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, Netherlands; ^bCentre for Science and Technology Studies, Leiden University, Leiden, Netherlands

ABSTRACT

We study regional patterns of scientific knowledge production in Europe using all scientific publications in the period 2000–2014 attributed to 813 scientific subfields. We show that the existing scientific portfolio of regions offers opportunities for related diversification and discourages the creation of knowledge on topics unrelated to the local knowledge base. Many lagging regions show clear growth, but complex knowledge production remains highly concentrated in regions in the North and West of Europe. For lagging regions there are advantages in not specializing too soon and to first diversify before moving into developing more complex knowledge.

ARTICLE HISTORY

Received 10 December 2018
Revised 5 July 2019
Accepted 12 July 2019

KEYWORDS

Smart specialization; indicators; scientific knowledge dynamics; path dependency; innovation policy

Introduction

We study regional patterns of scientific knowledge production in Europe using all scientific publications in the period 2000–2014 to understand the opportunities of different regions to develop new knowledge. It is now generally accepted that knowledge production is central to innovation and the economic performance of countries and regions (Arrow, 1962; Asheim et al., 2006; Romer, 1994). Fortunately, the potential for new scientific knowledge development is increasing rapidly. Specifically, the stock of available scientific knowledge is growing, there are many more researchers in more locations around the world than ever before (Grossetti et al., 2013; Maisonobe, Grossetti, Milard, Jégou, & Eckert, 2017), and more money is spent on research (UNESCO, 2010). Moreover, existing scientific developments apply positive feedback to knowledge production since a growing stock of successful ideas can be used as input to create new knowledge in the next stage (Arthur, 2007). Scientific knowledge is also more easily transferable across geographical space through the Internet (David & Foray, 2002; Heimeriks & Vasileiadou, 2008). New opportunities for increasingly complex knowledge developments are thus opening to benefit from the globally available stock of knowledge.

Nevertheless, knowledge production and accumulation remains very unevenly distributed over regions (Alkemade, Heimeriks, Schoen, Villard, & Laurens, 2015; Audretsch &

Feldman, 1996; Florida, 2005), and many regions struggle to replicate the levels of productivity and innovativeness in knowledge production achieved in leading regions (Heimeriks & Balland, 2016). In particular, highly specialized and complex outputs tend to be produced at relatively few locations and often provide long-run competitive advantage (Hausmann & Hidalgo, 2009; Hidalgo, Klinger, Barabasi, & Hausmann, 2007; Kogler, Heimeriks, Leydesdorff, Kogler, & Heimeriks, 2018). As a consequence, there is clear need to better understand the geographical distribution of knowledge production, as well as to better understand the potential of the places that ‘don’t matter’ (Rodríguez-Pose, 2018) in producing new scientific knowledge. Shaping the territorial dimension of future policies for sustainable growth requires understanding the knowledge development perspectives of different places. Each region has a unique perspective on global developments that emerged in path dependent process of specialization (Heimeriks & Boschma, 2014).

While it is clear that regions specialize in different fields of knowledge production, it remains an open question what fields and topics provide the best opportunities for further knowledge developments at different locations given the nature of the existing regional knowledge base (Heimeriks & Balland, 2016)? While some attention has been directed to processes of knowledge production from a spatial perspective using patent data, very little consideration has been given to different fields of knowledge produced within specific places or about relatedness of various knowledge domains in general, and within regions more specifically (Kogler, Rigby, & Tucker, 2013; Kogler 2015). Furthermore, it is unclear at how different scientific portfolios of regions enable regional specialization processes? Yet, these questions are vital if we are to make sensible policies towards science-driven development.

We address the question how scientific knowledge production is distributed over European regions and what strategies are available for lagging regions? We investigate the regional specialization patterns of scientific knowledge production in all European regions (i.e. NUTS2 regions) over the period 2000–2014. The aim of this study is to show how regional knowledge diversity and knowledge complexity shape patterns of regional knowledge production over time. This will provide a better understanding of regional opportunities and constraints for further scientific knowledge development and utilize each region’s unique knowledge portfolio to chart its path to development. A classification system of 813 scientific fields is used as an indication of scientific developments in all European regions over a period of time. While scientific publications only represent a part of the codified knowledge base of a region, they do provide a rich source of information about the local knowledge base that cannot be easily obtained from other sources, especially concerning knowledge developments that are not (yet) commercially exploited.

The theoretical starting point of this analysis is the idea that the dynamics of scientific knowledge are path and place dependent (Heimeriks & Boschma, 2014), and that the current research portfolio of a region influences the further capacity to produce knowledge. Scientific fields can be expected to constrain and facilitate the local opportunities of researchers to different degrees (Heimeriks & Balland, 2016). Some fields will be available to many regions, while other – more complex – fields require unique capabilities only available in top regions (Nomaler, Frenken, & Heimeriks, 2014). In addition to the amount of knowledge, we will thus focus on the diversity of knowledge that regions can produce, since this is expected to determine the ability of regions to move into closely related fields

of knowledge. We will also focus on the complexity of knowledge because it is expected to provide regions with long-lasting comparative advantages.

We adopt a portfolio-wide perspective in order to identify European regional opportunities and constraints to scientific knowledge developments. After presenting the theoretical context of this paper, we show how differences in regional scientific portfolios are operationalized. The results section presents for the first-time a regional analysis of patterns of European scientific knowledge production. Furthermore, we show how diversity, relatedness and knowledge complexity shape these patterns. The final sections provide conclusions and discuss some issues regarding smart specialization policy.

Theory

Increasingly, knowledge is considered an important source of economic growth (Arrow, 1962; Asheim et al., 2006; Romer, 1994). With the advent of the learning economy, it is widely agreed that economic growth is to a growing extent based on the production, appropriation and distribution of knowledge. Knowledge production is path dependent, which means that from an evolutionary perspective existing scientific knowledge provides the building blocks for further knowledge production (Arthur, 2007). Because knowledge builds atop knowledge, once learning starts it can be built dynamically and at scale (Arrow, 1962).

Knowledge production is also place dependent; it is differentiated among locations (Heimeriks & Boschma, 2014). As the overall stock of knowledge has expanded, a division of labour distributed its parts widely across different regions. The current scientific portfolio of a region influences the capacity to develop new ideas. New knowledge developments are dependent on place-specific circumstances that reflect conditions inherited from the previous knowledge production at a location (Martin & Sunley, 2007). Regions are thus increasingly understood as localized communities of practice that reflect place-bound sets of technological capabilities, routines and institutional arrangements (Storper, 1997). To capture both the limits and the potential new knowledge developments given an existing portfolio of knowledge, Kauffman coined the set of all possible new knowledge combinations ‘the adjacent possible’ (Kauffman, 1993). Learning processes are thus strongly shaped by contextual features (Antonelli, 2005). Those regions whose knowledge base and capabilities are closely related to new topics will be the locations which provide the greatest opportunities and possibilities for dynamic trajectories. In contrast, those locations whose knowledge base is locked-in to previous knowledge paradigms that are unrelated to new knowledge developments will find it very difficult to accumulate capabilities based on globally evolving topics and fields (Heimeriks & Boschma, 2014). Furthermore, knowledge developments are partially irreversible: once new topics and the accompanying skills and routines have moved on, previous or simpler topics are ‘forgotten’, and to reintroduce them would require a new learning process and the modification of individual and collective skills, organizational practices and institutions (Arthur, 1989).

We thus expect that the path dependent and place dependent nature of scientific knowledge production will result in concentration of activities within a limited number of scientific fields (Boschma, Heimeriks, & Balland, 2014; Heimeriks & Balland, 2016). As locations specialize in particular scientific topics, these offer opportunities for

further improvements in adjacent topics and fields, and discourage the creation of knowledge unrelated to the existing local knowledge base because of costs associated with search in that space (Heimeriks & Boschma, 2014; Leydesdorff, Heimeriks, & Rotolo, 2016; Rigby, 2013). The local accumulation of tacit knowledge provides an intangible asset that is difficult to copy by non-local agents, as geographical distance may form an barrier for the transfer of tacit knowledge (Breschi, Malerba, & Orsenigo, 2000; Gertler, 2003). The diversity of knowledge that regions can produce is thus expected to also determine the ability of regions to move into closely related fields of knowledge. Diversity matters because regions are more likely to expand and diversify into new topics and fields that are closely related to their existing activities. More diversity of existing topics provides more opportunities for diversification into related activities. Previous research has shown that the relatedness density between topics with the scientific activity of regions allows us to specify the ability of regions to diversify. The idea of relatedness density combines the information given by (i.e. the set of topics on which they publish (see Boschma et al., 2014 for a more technical description)).

Moreover, complexity of knowledge matters because it allows regions to produce idiosyncratic knowledge that few other regions can make. In analogy with the production of goods (Hausmann & Hidalgo, 2009; Hidalgo et al., 2007), it can be expected that sophisticated regions are capable of contributing to a large variety of 'exclusive' fields that few other regions can develop. Hausmann and Hidalgo (2009) developed a complexity measure based on the diversity of outputs within territories and the ubiquity (or range) of territories across which individual outputs can be produced. They argue that countries and regions develop different core competences: countries and regions that amass larger sets of capabilities tend to produce more specialized outputs that are hard to copy or imitate by others. The complexity of a region is embodied in the wide range of knowledge or capabilities that are combined to produce outputs: less ubiquitous outputs are more likely to require a greater variety of capabilities. We thus expect that these specialized (e.g. more complex) outputs tend to be produced at relatively few locations and often provide long-run competitive advantage.

The regional system of scientific research is evolving, leading to newer, more numerous, and increasingly dynamic nodes of scientific production throughout the globe (Grossetti et al., 2013). However, regions differ in the composition, as well as the quantity of knowledge production. Just as regions differ in size and wealth, they also vary in the diversity and complexity of their knowledge base. Especially large, metropolitan regions are capable of contributing to a wide range of fields (Nomaler et al., 2014). Previous research suggests that geographical skewness along urban-rural divide is growing even more pronounced with time (Balland & Rigby, 2017). Cities are increasingly the relevant platforms for knowledge production (Kogler et al., 2018). Truly innovative knowledge developments increasingly take place in a very select number of creative cities (Florida, Adler, & Mellander, 2017). This is visible in the pattern that many phenomena scale with city size and do this in a non-linear way. As city size increases, per capita quantities such as wages and GDP increase by approximately 15% more than the expected linear growth (Bettencourt, Lobo, Strumsky, & West, 2010). For knowledge production activities, scaling effects have been found to be even bigger (Nomaler et al., 2014). We thus expect that a high diversity of topics is increasingly found in large metropolitan areas.

The portfolio of scientific topics and fields can thus be expected to constrain and facilitate regional opportunities for knowledge production to different degrees. In particular, as a more complex knowledge base reflects a higher number of capabilities, one may expect that more complex regions will diversify more easily into new topics and fields. This poses an important challenge for regional policy-makers to decide how to make use of their limited investments across a range of scientific fields, especially in regions that do not possess a complex and diverse knowledge base. Catching up of these lagging regions is determined by their ability to absorb ideas and knowledge from the research frontier, which requires a long-term learning perspective (Hausmann & Rodrik, 2003). Policies in which governments intervene in the allocation of resources among fields of research can help lagging regions to improve and learn. Learning may be more marked in some complex fields than in others, and the benefits of that learning, including the institutional development required for success, may spill over to other fields (Arrow, 1962; Stiglitz, 1999).

Long term growth of regional knowledge and innovation thus requires pulling the regions' resources into those fields that allow for diversification and structural change (Rodrik, Subramanian, & Trebbi, 2004). The idea of 'Smart Specialisation' provides a rationale for regions to acquire new capabilities and create and mobilize institutions of knowledge production. 'Smart Specialisation' is an innovation policy concept intended to promote efficient and effective use of public investment in research. Its goal is to boost regional innovation by enabling regions and cities to focus on their strengths (Foray, David, & Hall, 2009). Smart specialization means identifying the unique innovative characteristics and assets of each region, highlighting each region's competitive advantages (McCann & Ortega-Argilés, 2013).

This presents a knowledge gap. Shaping the territorial dimension of future policies for knowledge-based growth requires understanding the territorial diversity of different places. Each region has a unique perspective on global developments (Heimeriks & Boschma, 2014). Some regions have a diverse knowledge base, while many others are capable of producing new knowledge in only a small number of fields. Furthermore, some regions produce ubiquitous knowledge while others produce more specialized and complex topics that are hard to copy or imitate by others. We expect that pursuing complex knowledge is a high risk strategy for lagging regions that is likely to fail. We thus hypothesize that for lagging regions, the best strategy is to diversify into adjacent fields thus exponentially increasing the potential for new combinations and opportunities for further diversification.

Following from this discussion of the relevant literature, we expect that regions can be characterized by distinct scientific portfolios. From a policy perspective, portfolio analyses inform policy-makers in their mission to make best use of the existing scientific strengths of regions. To the best of our knowledge, the results will for the first time provide the opportunity for the comparison of European regional scientific knowledge portfolio along various measures and dimensions.

Data and methods

In this study, we use scientific journal publications as an indicator for regional knowledge production. We focus on the portfolio of topics in European regions. Adoption of a

portfolio-wide perspective is necessary for understanding and managing research developments in order to identify a region's opportunities and constraints to knowledge development and utilize that region's own unique knowledge portfolio to chart its path to prosperity (Leydesdorff et al., 2016; Wallace & Rafols, 2015). In order to show how the knowledge base of regions affects the development of new knowledge, we focus on lagging regions with less than one thousand publications in the period 2000–2002.

Publication data over the period 2000–2014 are retrieved from the Web of Science. Based on the country and city indicated in the author affiliations, we were able to allocate scientific publications to NUTS-2 level regions for all EU-28 countries plus Iceland, Norway and Switzerland. In order to identify the specific research areas in which the regions are active, we relied on a classification system developed at CWTS (Waltman & Van Eck, 2012). This system is based on citation relations between publications, i.e. publications that cite each other more than others are clustered together. Compared to other frequently used classification systems (e.g. the WoS Subject Categories, that classify publications according to the journals where they were published), this classification is much more nuanced as it is built at the level of individual publications and is not dependent on pre-defined categories. It also helps to overcome the problem of classifying papers published in 'multidisciplinary journals'. All publications in our dataset are uniquely attributed to a subfield. In this study, we use the intermediate level of 813 subfields in the CWTS classification system.

Relatedness density

In order to understand the ability of regions to diversify into related subfields, we first need to measure the scientific relatedness among subfields. We use the region-subfields network approach following (Hidalgo et al., 2007). The relatedness between scientific subfield i and scientific subfield j $\varphi_{i,j,t}$ is computed by taking the minimum of the pair-wise conditional probabilities of regions published in one scientific subfield i , given that they published in another scientific subfield j during the same period. To avoid the noise induced by negligible patenting activity, we only consider the scientific subfields in which regions have a revealed comparative advantage (RCA):

$$\varphi_{ijt} = \min(P(RCA_{it}|RCA_{jt}), P(RCA_{jt}|RCA_{it}))$$

where RCA is a binary variable that assumes the value 1 (positive specialization) when a region produced a greater share of publications in subfields i than EU-28 plus Iceland, Norway and Switzerland as a whole; and 0 otherwise.

$$\frac{\text{Publication}_{rit} / \sum_i \text{Publication}_{rit}}{\sum_r \text{Publication}_{rit} / \sum_r \sum_i \text{Publication}_{rit}} > 1$$

We compute relatedness $\varphi_{i,j,t}$ between each pair of scientific subfields i and j for five different non-overlapping periods: 2000–2002, 2003–2005, 2006–2008, 2009–2011 and 2012–2014.

In the next step, we create a region-subfield level variable relatedness density that combines the information given by the relatedness between subfields with the scientific activity of regions (i.e. the set of subfields in which they publish (see Boschma et al., 2014 for a more technical description)). The relatedness density measure thus helps us to capture

how much of the scientific subfield produced within each NUTS-2 region tends to build on the existing scientific knowledge base of regions. We calculated the ‘relatedness density’ index following Hidalgo et al. (2007) and Boschma, Balland, and Kogler (2015).

$$\text{Relatedness}_{i,r,t} = \frac{\sum_{j \in r, j \neq i} \varphi_{ijt}}{\sum_{j \in r} \varphi_{ijt}} * 100$$

The relatedness around a subfield i in a region r at period t is the sum of scientific relatedness φ_{ijt} of subfield i to all other subfields in which the region specializes in ($\text{RCA}_{rit} > 1$), divided by the sum of scientific relatedness of subfield i to all other subfields at time t .

Scientific complexity

In order to capture the complexity of scientific subfields, we follow (Hausmann & Hidalgo, 2009) and (Balland & Rigby, 2017) by analysing the structure of the region-subfields network. A knowledge complexity index needs to combine information on both the 2-mode degree distribution of regions (diversity) and the 2-mode degree distribution of the topics it produces (ubiquity). Therefore, we follow Hausmann and Hidalgo (2009) and sequentially combine the diversity of regions and ubiquity of topics. The complexity index reflects the difficulty of mastering the capabilities required to produce knowledge in a particular scientific subfield, the diversity of capabilities within regions and the relatedness between them. We computed the knowledge complexity index using the EconGeo package in R developed by (Balland, 2017).

Regional scientific diversification

In the next step, we want to estimate how relatedness density and knowledge complexity influence the process of **entry** of scientific subfields in regions’ portfolios. We model knowledge dynamics as the process of entry of scientific topics in regional portfolios (i.e. as an evolving region–topic network). This will allow us to test the hypothesis that the ability of regions to move into new fields of knowledge is determined by the relatedness of new fields to the existing portfolio of fields of a region. We regress the emergence of new scientific subfields on their degree of relatedness with the scientific portfolio of regions (which is captured by the relatedness density), and their degree of complexity. The econometric equations to be estimated can be written as follows:

$$\text{Entry}_{i,r,t} = \beta_0 + \beta_1 \text{Relatedness}_{i,r,t-1} + \beta_2 \text{Complexity}_{i,t-1} + \beta_3 \text{Complexity}_{i,t-1} \\ * \text{Relatedness}_{i,r,t-1} + \beta_4 \text{Region}_{r,t-1} + \beta_5 \text{Subfields}_{i,t-1} + \varphi_r + \alpha_t + \varepsilon_{i,r,t}$$

where $\text{Region}_{r,t-1}$ and $\text{Subfields}_{i,t-1}$ are the natural logarithm of their respective total number of publications in region r and subfield i at time $t-1$. Tables 1 and 2 provide the summary statistics and correlation statistics for the variables. Since the number of publications in a region highly correlates with relatedness density, we exclude this control variable in the regression.

We estimated the entry model using a linear probability model (LPM) to assess the probability that a region develops a new specialization in a scientific subfield. The baseline specification is a two-way fixed effects model where φ_r is a region fixed effect, α_t is

Table 1. Summary statistics.

Statistic	<i>N</i>	Mean	St. Dev.	Min	Max
Entry	680,426	0.126	0.331	0	1
Relatedness density	680,426	25.557	9.782	0.120	51.650
Size region	680,426	7.579	1.731	1.099	11.526
Scientific complexity	680,426	50.871	19.635	0.000	100.000
Size subfields	680,426	7.262	0.923	4.060	10.192

Table 2. Correlation statistics.

	Relatedness density	Size region	Scientific complexity	Size subfields
Relatedness density	1.000	0.931	0.015	0.051
Size region	0.931	1.000	0.036	0.042
Scientific complexity	0.015	0.036	1.000	0.038
Size subfields	0.051	0.042	0.038	1.000

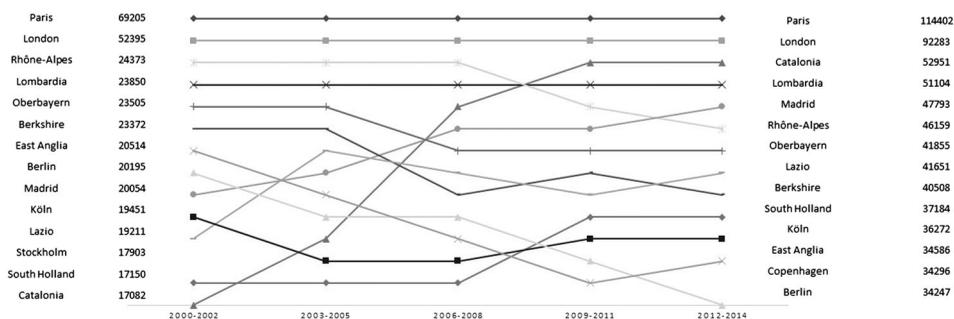
a time fixed effect and $\varepsilon_{i,r,t}$ is a regression residual. We cluster the standard errors for all regression outputs at the region and subfields level. Our panel consists of data for 286 NUTS-2 regions and 813 scientific subfields over the period 2000–2014. We average the data over non-overlapping three-year periods, denoted by t . All independent variables are lagged by one period to avoid potential endogeneity issues, denoted by $t-1$.

Results

Descriptives

The regional landscape in scientific knowledge production is very spiky. There are significant differences among regions in their capacity to produce science. Most regions produce a modest amount of publications (less than 10k per annum), while a few regions are very productive. The most productive regions are London, Paris, Catalonia, Rhones-Alpes and Lombardia. These regions are all within the EU15 (see Figure 1).

Two Spanish regions, Catalonia and Madrid moved up the ranking considerably since 2000, while Rhones-Alpes, Berkshire, East Anglia, and Berlin show a relative decline. However, all the most important regions substantially increased their publication

**Figure 1.** Ranking of the most important knowledge producing regions in Europe (2000–2002, 2006–2008, 2011–2014).

output in the period under study. Paris and London remain the knowledge production powerhouses of Europe (Figure 2).

As expected, the scientific portfolio of most regions comprises only a small subset of the range of possibilities that define knowledge space. When we normalize the production of publications against the number of inhabitants, the strong position of North-west Europe emerges, but the uneven distribution of scientific knowledge production over regions remains. The number of publications per capita clearly shows that regions in North-west Europe (U.K., Belgium, The Netherlands, Denmark, Sweden and Finland) are clearly outperforming the other European regions. It is noteworthy that German regions are characterized by a relatively low per capita scientific knowledge production.

Clearly, there is a serious divide between regions in terms of their capacity for knowledge production. The question is whether convergence among the regions is taking place, i.e. whether lagging regions are able to closing the gap. Figure 3 shows the annual growth rate in scientific knowledge production of European regions.

Only the U.K. region of Cumbria shows a clear decline in scientific output, while all other regions show growth. Most pronounced growth takes place in Southern Europe (Portugal, Spain and Greece) and in Central and Eastern Europe, especially in Romania.

Diversity

The most productive regions are generally also the most drivers in terms of the number of topics covered. This further confirms the notion that metropolitan areas are increasingly important drivers of our knowledge-based economy (Florida, 2002). In addition to Paris and London, we notice again an impressive rise of the two Spanish metropolitan regions of Madrid and Barcelona (Figure 4).

Large capital regions in the EU15 dominate the diversity ranking. Because density spurs knowledge production and innovation by bringing people and ideas together and enabling them to combine and recombine in new ways, large metropolitan regions with their dense mixtures of people, universities, research organizations and companies are increasingly prominent locations of knowledge production. In other words, proximity increases the circulation not only of goods and people, but of ideas as well (Nomaler et al., 2014). As a consequence, especially large cities can be expected to benefit from the diversity of human resources and institutional resources to yield greater output in terms of technological developments (Bettencourt, Lobo, Helbing, Kühnert, & West, 2007).

Every region has its own, unique knowledge base. Regions specialize because of the cumulative and path-dependent character of scientific knowledge production (Heimeriks & Boschma, 2014). The opportunities to diversify into new fields is to a large extent dependent on the existing portfolio of related knowledge (i.e. the adjacent possible). From this perspective it is clear that the diversity of the knowledge base can be considered an important indicator for further knowledge developments in regions. Portfolio analysis helps us to understand the capabilities that make up a regions portfolio. After all, almost all regions (except the simplest and the smallest) are involved in more than one subfield.

Following (Hausmann & Hidalgo, 2011), the scientific complexity provides a measure of the production characteristics of European regions. As most of the measurements used in complexity economics, the goal of this index is to explain the quality of a regional

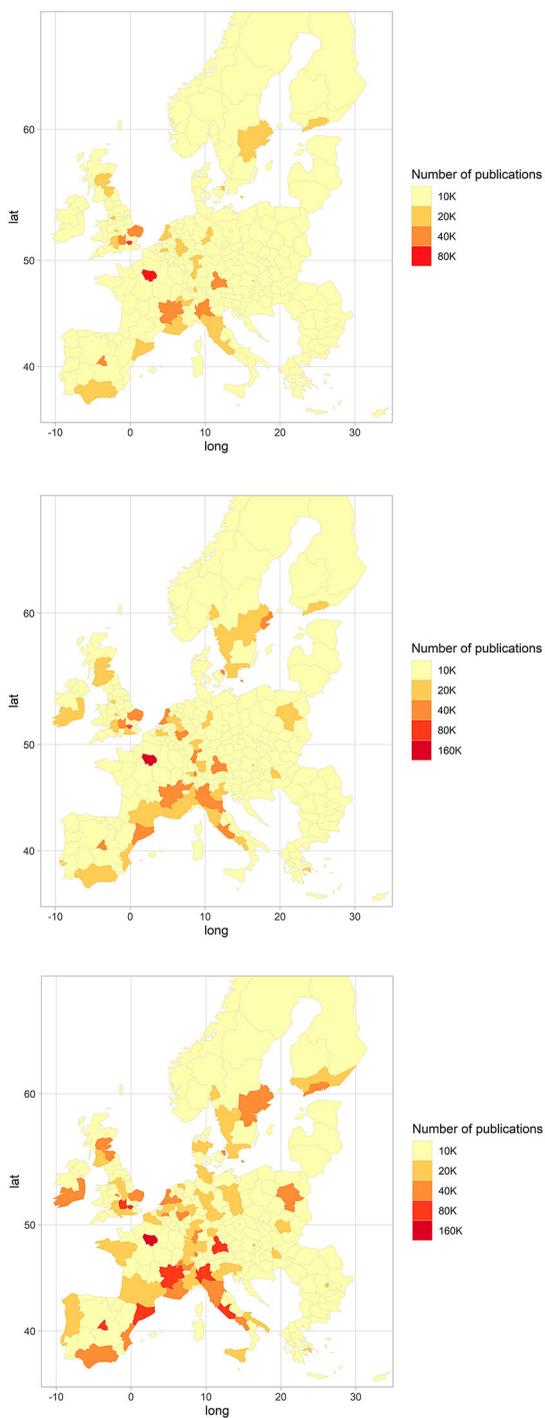


Figure 2. Distribution of publications over NUTS-2 regions (2000–2002, 2006–2008, 2011–2014).

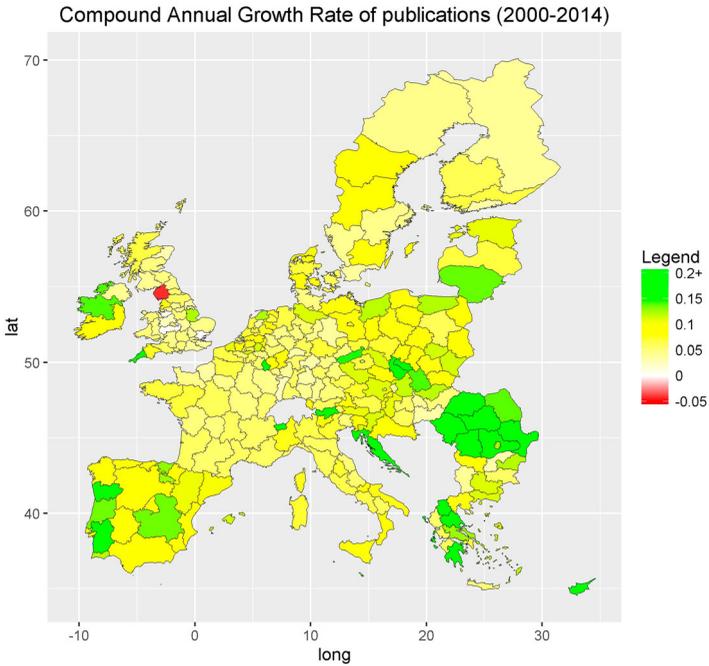


Figure 3. Growth rate in publication output of European regions between 2000 and 2014.

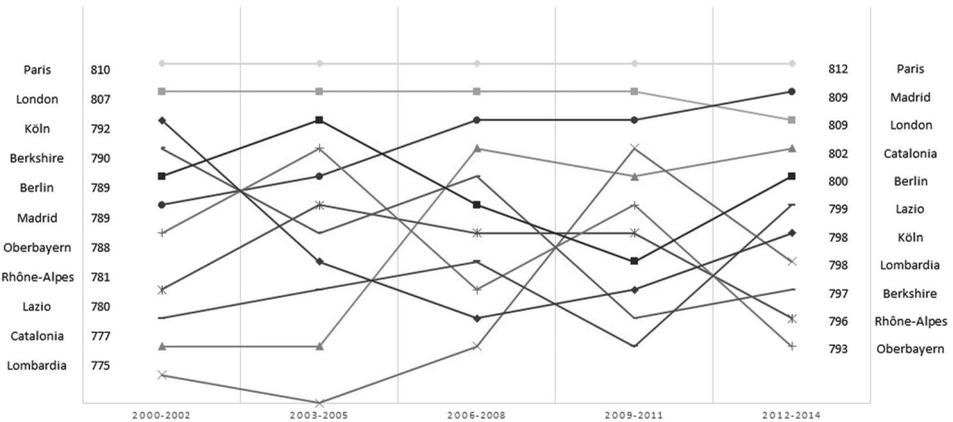


Figure 4. Ranking of the most diverse knowledge producing regions in Europe as indicated by their activity in scientific subfields (2000-2002, 2006-2008, 2011-2014).

knowledge base as a whole rather than the sum of its parts. The complexity measure looks to explain the knowledge produced in a region combining metrics of the diversity of regions and the ubiquity of the fields to create measures of the relative complexity of a regions scientific portfolio (Figure 5).

The regions located in the North-West of Europe are those with the most complex scientific output on average. Central and Eastern European regions are characterized by relatively low levels of complex knowledge production.

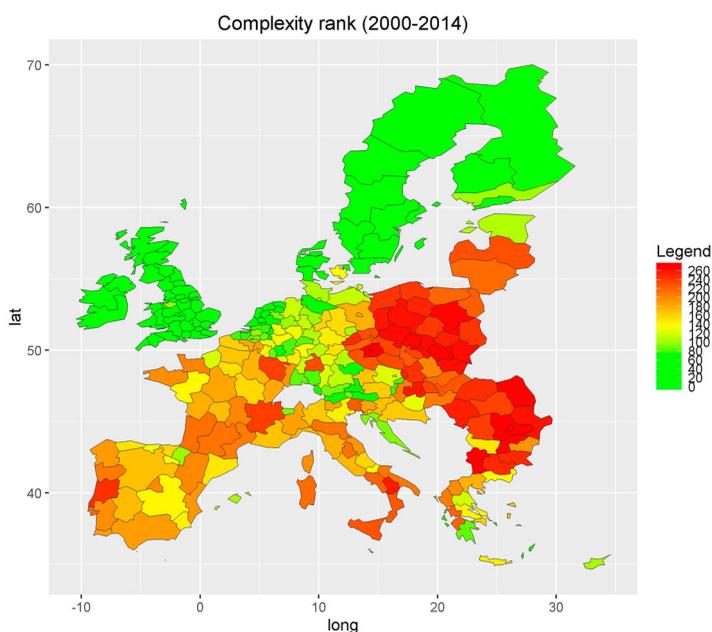


Figure 5. The average complexity of the scientific knowledgebase of European NUTS region.

Entry model

We showed that the scientific portfolio of most regions comprises only a small subset of the range of fields in the total knowledge space. In the next step, we analyse whether scientific fields within a regional portfolio indeed offer opportunities for further improvements in adjacent fields, and discourage the creation of knowledge unrelated to the existing local knowledge base.

We use the relatedness density measure to capture how much of the scientific subfield produced within each region tends to build on the existing scientific knowledge base of regions. The results show that relatedness density has a positive and significant effect on the probability that a region specializes ($RCA > 1$) in a new scientific subfield in all entry models in Table 3, which is consistent with other findings (Heimeriks & Balland, 2016; Heimeriks and Boschma, 2014). The effect of relatedness density is also strong: an increase of 10% in relatedness is associated with a 28% relative increase in the probability of entry. The effects of scientific complexity on entry are positive and significant in model (3) and (4). However, the effects of scientific complexity on entry turn negative in model (5) and (6) after adding the interacting term of relatedness and complexity. The coefficients of the interacting term are significantly positive, indicating that regions are more likely to enter complex scientific subfields related to their existing scientific knowledge base. Because the diversity of topics indicates a diversity of capabilities, the results confirm the idea that more diverse capabilities are important for producing complex knowledge.

Entry model for different types of region

To further investigate the place-dependent process, we split the sample between regions with different levels of overall scientific complexity and diversity (see Figure 6).

Table 3. Entry model – full sample.

	Dependent variable: entry (=1) 2000–2014					
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.176*** (0.016)	0.104*** (0.015)	0.023*** (0.005)	0.087*** (0.015)	0.123*** (0.019)	0.089*** (0.017)
Relatedness Density		0.003*** (0.002)	0.003*** (0.0002)	0.003*** (0.0002)	0.002*** (0.0003)	0.017*** (0.001)
Scientific Complexity			0.0004*** (0.0001)	0.0004*** (0.0001)	-0.0003** (0.0002)	-0.0004** (0.0002)
Relatedness Density* Scientific Complexity					0.00003*** (0.00001)	0.00003*** (0.00001)
Size Subfields	-0.007*** (0.002)	-0.009*** (0.002)		-0.009*** (0.002)	-0.009*** (0.002)	-0.011*** (0.002)
Region fixed effects	No	No	No	No	No	Yes
Time fixed effects	No	No	No	No	No	Yes
Observations	680,426	680,426	680,426	680,426	680,426	680,426
R ²	0.0004	0.010	0.010	0.011	0.011	0.025
Adjusted R ²	0.0004	0.010	0.010	0.011	0.011	0.025

Notes: The dependent variable entry equals 1 if a region *r* gains new relative comparative advantage in a given scientific subfield *l* during the corresponding three-year window; 0 otherwise. All independent variables are mean centred and lagged by one period. Heteroskedasticity-robust standard errors (clustered at the region and scientific subfield level) are shown in parentheses.

Coefficients are statistically significant at the **p* < .1; ***p* < .05; ****p* < .01.

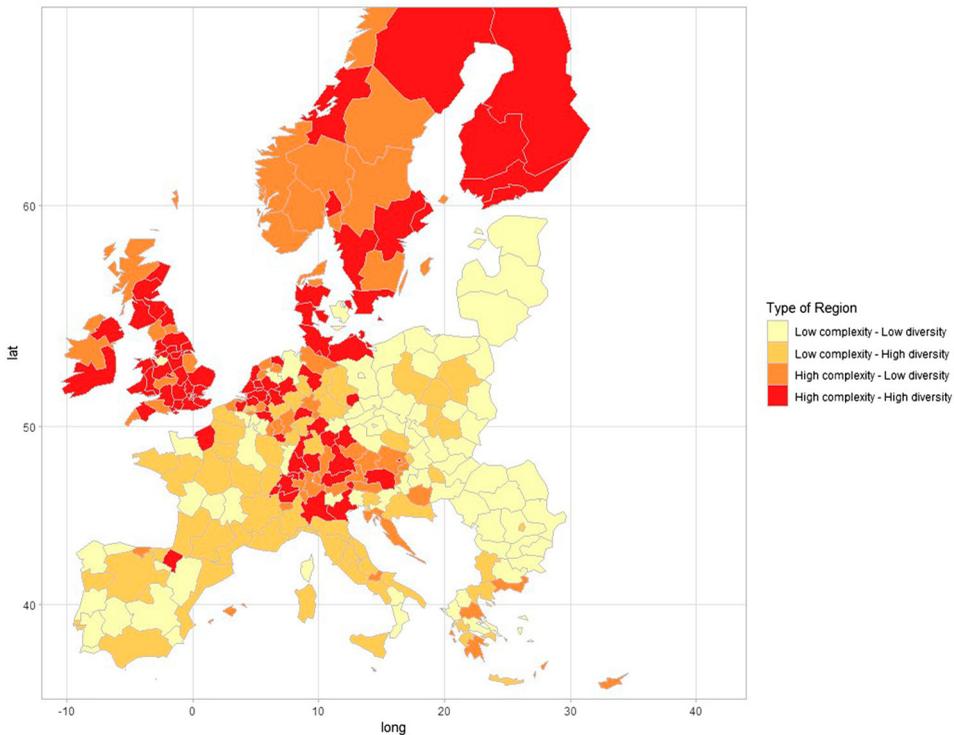


Figure 6. Type of NUTS-2 regions based on overall scientific complexity and diversity (2000–2002).

We use the median value of knowledge complexity and diversity at each period to categorize regions into 4 types (Low complexity – Low diversity, Low complexity – High diversity, High complexity – low diversity and High complexity – High diversity). As

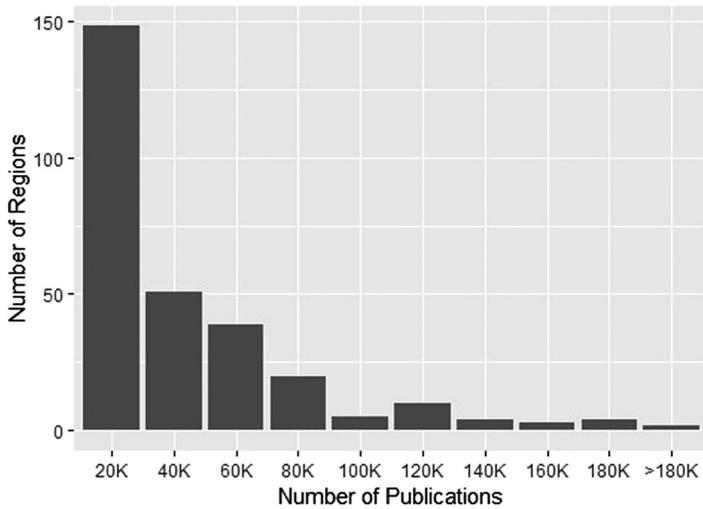


Figure 7. Distribution of regions in number of publications between 2000 and 2014.

shown before, complex knowledge production is very concentrated in the north-western parts of Europe. Other regions in north western Europe can be characterized as low diversity-high complexity. Regions in France, Spain and Italy are mostly producing diverse knowledge of low complexity. Regions in central and eastern Europe are mostly procuring knowledge of low diversity and low complexity.

Figure 7 shows the distribution of regions in terms of the number of publications during the period 2000–2014. The distribution is skewed and most regions have less than 60,000 publications in the period under study.

In order to show how the knowledge base of regions affects the development of new knowledge, we focus on lagging regions with less than one thousand publications in the period 2000–2002. There are 99 regions in this category, accounting for 34% of all regions. Figure 8 shows that there is a significantly positive correlation between the diversity of regional knowledge base and the number of new topics. Regions with more diverse knowledge base can develop more new topics. Furthermore, Figure 9 shows that there is also a significantly positive correlation between the diversity of regional knowledge base and the average complexity of new topics. The diversity of regional knowledge base can also help develop more complex topics.

Entry of new topics clearly shows different patterns among those four types of regions. The results of the entry model are presented in Tables 4 and 5. The coefficients of relatedness density are significant and positive in almost all models of entry. However, the coefficients of scientific complexity are significantly negative in regions with low overall complexity level (Table 4), significantly positive in regions with high overall complexity level (Table 5). The coefficients of the interacting term are significantly positive in all types of region except regions with low complexity and low diversity. However, the coefficients of the interacting term are larger in regions with high knowledge diversity.

In sum, our findings suggest the overall scientific complexity of regions is important for developing new knowledge in complex scientific subfields. Regions with diverse scientific knowledge base can overcome the complex dilemma by developing complex scientific

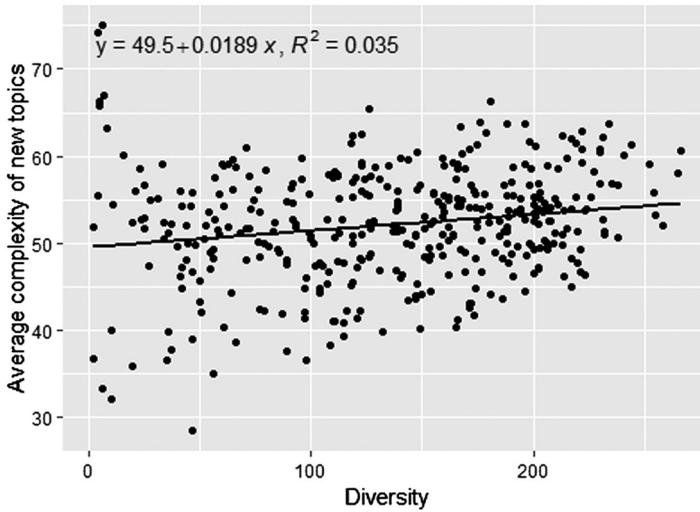


Figure 9. Correlation between diversity of regional knowledge base and the average complexity of new topics in lagging regions.

subfields related to their existing knowledge base. For regions with both low level of complexity and low level of diversity, they should focus on scientific subfields related to their existing scientific knowledge base to develop capabilities for future diversification into complex subfields.

Discussion

This study focused on the dynamics of knowledge as made visible by scientific journal publications. From a smart specialization perspective (Foray, Mowery, & Nelson, 2012),

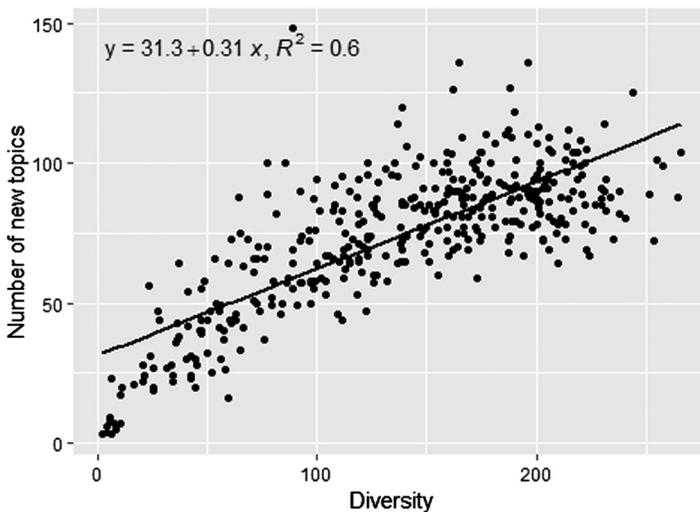


Figure 8. Correlation between diversity of regional knowledge base and the number of new topics in lagging regions.

Table 4. Entry model by types of region.

	Dependent variable: Entry (=1) 2000–2014					
	Low complexity – low diversity			Low complexity – high diversity		
	(1)	(2)	(3)	(1)	(2)	(3)
Constant	0.059*** (0.019)	0.052** (0.024)	0.096*** (0.033)	0.130*** (0.033)	0.371*** (0.063)	−0.813*** (0.093)
Relatedness Density	0.004*** (0.000)	0.005*** (0.001)	0.020*** (0.002)	0.007*** (0.001)	−0.0003 (0.002)	0.034*** (0.003)
Scientific Complexity	−0.001*** (0.000)	−0.001*** (0.000)	−0.001*** (0.000)	−0.001*** (0.000)	−0.005*** (0.001)	−0.002* (0.001)
Relatedness Density* Scientific Complexity		−0.00001 (0.000)	0.00002 (0.000)		0.0001*** (0.000)	0.0001*** (0.000)
Size Subfields	0.003 (0.003)	0.003 (0.003)	−0.0004 (0.003)	−0.026*** (0.002)	−0.026*** (0.002)	−0.029*** (0.002)
Region fixed effects	No	No	Yes	No	No	Yes
Time fixed effects	No	No	Yes	No	No	Yes
Observations	200,646	200,646	200,646	141,286	141,286	141,286
R ²	0.015	0.015	0.027	0.011	0.011	0.031
Adjusted R ²	0.015	0.015	0.026	0.011	0.011	0.03

Notes: The dependent variable entry equals 1 if a region *r* gains new relative comparative advantage in a given scientific subfield *l* during the corresponding three-year window; 0 otherwise. All independent variables are mean centred and lagged by one period. Heteroskedasticity-robust standard errors (clustered at the region and scientific subfield level) are shown in parentheses.

Coefficients are statistically significant at the * $p < .1$; ** $p < .05$; *** $p < .01$.

there are good reasons to focus on the localized production and accumulation of scientific knowledge (Heimeriks & Balland, 2016). Scientific knowledge production provides unique insights in the regional capabilities. While economic opportunities are relatively invariant across different regions (Breschi, Lissoni, & Malerba, 2003), knowledge bases are more likely to differ according to their geographical locations. Indeed, it has been shown that the knowledge production and accumulation are more geographically concentrated

Table 5. Entry model by types of region.

	Dependent variable: Entry (=1) 2000–2014					
	High complexity – low diversity			High complexity – high diversity		
	(1)	(2)	(3)	(1)	(2)	(3)
Constant	−0.095*** (0.021)	−0.065*** (0.022)	−0.096*** (0.020)	0.021 (0.029)	0.262*** (0.066)	−0.746*** (0.093)
Relatedness Density	0.005*** (0.000)	0.003*** (0.001)	0.027*** (0.003)	0.007*** (0.001)	−0.0003 (0.002)	0.026*** (0.003)
Scientific Complexity	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.002*** (0.000)	−0.003*** (0.001)	−0.003*** (0.001)
Relatedness Density* Scientific Complexity		0.00003*** (0.000)	0.00002* (0.000)		0.0002*** (0.000)	0.0001*** (0.000)
Size Subfields	0.007** (0.003)	0.007** (0.003)	0.005* (0.003)	−0.027*** (0.002)	−0.027*** (0.002)	−0.030*** (0.002)
Region fixed effects	No	No	Yes	No	No	Yes
Time fixed effects	No	No	Yes	No	No	Yes
Observations	170,320	170,320	170,320	168,174	168,174	168,174
R ²	0.029	0.03	0.046	0.021	0.022	0.038
Adjusted R ²	0.029	0.03	0.045	0.021	0.022	0.037

Notes: The dependent variable entry equals 1 if a region *r* gains new relative comparative advantage in a given scientific subfield *l* during the corresponding three-year window; 0 otherwise. All independent variables are mean centred and lagged by one period. Heteroskedasticity-robust standard errors (clustered at the region and scientific subfield level) are shown in parentheses.

Coefficients are statistically significant at the * $p < .1$; ** $p < .05$; *** $p < .01$.

than economic activities (Florida, 2005). The unique innovative potential of regions is thus strongly linked to their ability to develop an institutional context that facilitates the production and the accumulation of knowledge. The geographical patterns found here in relation to different evolutionary patterns of the global knowledge base, are consistent with earlier findings that market developments across sectors are largely determined by the level of accumulateness of the knowledge base (Malerba & Orsenigo, 2002).

We showed that relatedness clearly has a positive and significant effect on the probability that a new field enters the scientific portfolio of a region. The analysis shows a pronounced pattern of path- and place-dependence. However, some fields are available to many regions, while other fields require unique capabilities only available in a few scientifically advanced regions. An important finding here is that fields differ greatly in the opportunities they provide for further diversification. In general, regions in the North and West of Europe are on average most active in more complex topics of knowledge production. Furthermore, the patterns of knowledge production remain fairly stable over time.

It is clear that both diversity and complexity matter for knowledge development. The ability to produce complex knowledge is expected to be more valuable for regions compared to ubiquitous knowledge (Balland & Rigby, 2017). Scientific complexity of regions is shown to be important for developing new knowledge in complex scientific subfields. Regions with a diverse scientific knowledge base are most likely to develop more complex knowledge when these complex scientific subfields are related to their existing knowledge base. For regions with both low level of complexity and low level of diversity, the best strategy is to diversify by focusing on scientific subfields related to their existing scientific knowledge base in order to develop capabilities for future diversification into more complex subfields. Thus, in line with (Barzotto, Fai, & Tomlinson, 2018), the results suggest that for lagging regions there may be advantages in not specializing too soon and to diversify before moving into developing more complex knowledge.

A specialization process towards complexity may thus preferably start when regions have sufficiently developed diverse knowledge base. When the path dependent local knowledge base is narrow and offers few new opportunities for growth, regional developments may benefit from engaging in building new complementarities and synergies with different types of knowledge in a process of that may lead to new trajectories (Castaldi, Frenken, & Los, 2015).

Conclusion

We study regional patterns of scientific knowledge production in Europe using all publications from 813 scientific subfields in the period 2000–2014. For the first time, we show the diversity in regional scientific knowledge production. In general, we see pronounced patterns of specialization, each region is characterized by a unique portfolio of topics and fields.

The distribution of publications over regions is very skewed. Most of the regions produce a modest amount of publications (less than 10k per year), while a few regions are very productive. Especially large metropolitan areas benefit from the diversity of human resources and institutional complementarities that provide comparative

advantages to yield greater output in terms of knowledge production (Balland, Boschma, Crespo, & Rigby, 2018; Bonaccorsi & Thoma, 2007; Nomaler et al., 2014).

Scientific knowledge production is growing rapidly in Europe in almost all regions. Concerning pattern of catching up, we see that the relative growth of publication output is especially strong in the lagging regions. However, the gap between leading and lagging regions remains. Especially the ability to develop 'complex knowledge' remains concentrated in the advanced regions of north-western Europe. Convergence in levels of knowledge production among regions is anything but automatic (Hausmann & Rodrik, 2003). It is dependent on specific policies and institutional arrangements that have proved hard to identify and implement. Indeed, the recipes seem to vary from context to context (Tödting & Tripl, 2005). The divide between the research and innovation leaders in the north of Europe and the innovation laggards is proving very persistent. The process of convergence, which was already very slow before 2008, has since stalled, and has more recently moved into reverse (Veugelers, 2016). Catching up in Central and Eastern European countries is concentrated in urban metropolitan areas. Regions that narrowed the knowledge production gaps tend to benefit from economically vibrant locations in well-functioning cities.

Regarding the question what fields and topics provide the best opportunities for further knowledge developments at different locations given the nature of the existing regional knowledge base we showed that regional scientific knowledge production is characterized by path and place dependency. The existing scientific portfolio of regions offers opportunities for related diversification and discourages the creation of knowledge on topics unrelated to the local knowledge base. The entry of new fields in the regional knowledge portfolio is strongly correlated with their degree of relatedness with the existing scientific portfolio of regions. The effect of relatedness density is very pronounced, an increase of 10% in relatedness is associated with a 28% relative increase in the probability of entry to new fields of research.

Exploring the regional specialization in scientific publications that regions produce over time, we investigated the nature of the scientific portfolio that characterize the leading and lagging regions. Some fields are available to many regions, while other fields require unique capabilities only available in a few advanced regions. Specifically, the results show that regions are more likely to enter complex scientific subfields related to their existing scientific knowledge base. The overall scientific complexity of regions is important for developing new knowledge in complex scientific subfields. Regions with diverse scientific knowledge base can overcome the complexity dilemma by developing complex scientific subfields related to their existing knowledge base. Regions with both low level of complexity and low level of diversity, better focus on scientific subfields related to their existing scientific knowledge base to develop capabilities for future diversification into complex subfields.

From a policy perspective, the smart specialization rationale in which governments intervene in the allocation of resources among fields of research is shown to be helpful for lagging regions to improve and learn. Starting from the unique characteristics of the scientific knowledge base of each region, the results of this study show that related diversification is the best development strategy. For lagging regions there are clear advantages in not specializing too soon and to first diversify and increase the option for further knowledge developments before moving into developing more complex knowledge.

Acknowledgements

The authors would like to thank Dieter Kogler, Peter Kedron and Koen Frenken for their helpful suggestions.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Gaston Heimeriks  <http://orcid.org/0000-0002-0577-6938>

References

- Alkemade, F., Heimeriks, G., Schoen, A., Villard, L., & Laurens, P. (2015). Tracking the internationalization of multinational corporate inventive activity: National and sectoral characteristics. *Research Policy*, 44, 1763–1772. doi:10.1016/j.respol.2015.01.007
- Antonelli, C. (2005). Models of knowledge and systems of governance. *Journal of Institutional Economics*, 1, 51–73. doi:10.1017/S1744137405000044
- Arrow, K. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, 29, 155–173.
- Arthur, B. W. (1989). Competing technologies, increasing returns, and lock-in by historical events. *The Economic Journal*, 99, 116–131. doi:10.2307/2234208
- Arthur, W. B. (2007). *The nature of technology: What it is and how it evolves*. New York: The Free Press.
- Asheim, B., Boschma, R. A., Cooke, P., Dahlstrand-Lindholm, A., Laredo, P., & Piccauga, A. (2006). Constructing regional advantage. Principles, perspectives, policies. (D. G. Research, Ed.). European Commission. Retrieved from <http://lup.lub.lu.se/record/617551>
- Audretsch, D. B., & Feldman, M. P. (1996). R&D Spillovers and the geography of innovation and production. *The American Economic Review*, 86, 630–640. Retrieved from <http://www.jstor.org/stable/2118216>
- Balland, P. A. (2017). *Economic geography in R: Introduction to the EconGeo package*. Ssrn 3. doi:10.2139/ssrn.2962146
- Balland, P. A., Boschma, R., Crespo, J., & Rigby, D. L. (2018). Smart specialization policy in the European Union: Relatedness, knowledge complexity and regional diversification. *Regional Studies*, 1–17. doi:10.1080/00343404.2018.1437900
- Balland, P. A., & Rigby, D. (2017). The geography of complex knowledge. *Journal of Economic Geography*, 93, 1–23. doi:10.1080/00130095.2016.1205947
- Barzotto, M., Fai, F., & Tomlinson, P. R. (2018). Enhancing innovative capabilities in lagging regions: An extra-regional collaborative approach to RIS3. 2018 SMARTER Conference. Regional Studies Association, Seville, pp. 1–26.
- Bettencourt, L. M. A., Lobo, J., Helbing, D., Kühnert, C., & West, G. B. (2007). Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences of the United States of America*, 104(17), 7301–7306. doi:10.1073/pnas.0610172104
- Bettencourt, L. M. A., Lobo, J., Strumsky, D., & West, G. B. (2010). Urban scaling and its deviations: Revealing the structure of wealth, innovation and crime across cities. *PLoS One*, 5, e13541. doi:10.1371/journal.pone.0013541
- Bonaccorsi, A., & Thoma, G. (2007). Institutional complementarity and inventive performance in nano science and technology. *Research Policy*, 36, 813–831. doi:10.1016/j.respol.2007.02.009
- Boschma, R., Balland, P.-A., & 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, 223–250. doi:10.1093/icc/dtu012

- Boschma, R., Heimeriks, G., & Balland, P.-A. P.-A. (2014). Scientific knowledge dynamics and relatedness in biotech cities. *Research Policy*, 43, 107–114. doi:10.1016/j.respol.2013.07.009
- Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. *Research Policy*, 32, 69–87. doi:10.1016/s0048-7333(02)00004-5
- Breschi, S., Malerba, F., & Orsenigo, L. (2000). Technological regimes and schumpeterian patterns of innovation. *The Economic Journal*, 110, 388–410. doi:10.1111/1468-0297.00530
- Castaldi, C., Frenken, K., & Los, B. (2015). Related variety, unrelated variety and technological breakthroughs: An analysis of US state-level patenting. *Regional Studies*, 49, 767–781. doi:10.1080/00343404.2014.940305
- David, P. A., & Foray, D. (2002). An introduction to economy of the knowledge society. *International Social Science Journal*, 54, 9–23. doi:10.1111/1468-2451.00355
- Florida, R. (2002). *The rise of the creative class. And how it's transforming work, leisure and everyday life*. New York: Basic Books.
- Florida, R. (2005, October). The world is spiky. *The Atlantic Monthly*, pp. 48–51.
- Florida, R., Adler, P., & Mellander, C. (2017). The city as innovation machine. *Regional Studies*, 51(1), 86–96. doi:10.1080/00343404.2016.1255324
- Foray, D., David, P., & Hall, B. (2009). *Smart specialisation – The concept*. (Knowledge Economists Policy Brief No. 9, June). Brussels: European Commission, DG Research.
- Foray, D., Mowery, D. C., & Nelson, R. R. (2012). Public R&D and social challenges: What lessons from mission R&D programs? *Research Policy*, 41(10), 1697–1702. doi:10.1016/j.respol.2012.07.011
- Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of Economic Geography*, 3, 75–99. doi:10.1093/jeg/3.1.75
- Grossetti, M., Eckert, D., Gingras, Y., Jegou, L., Lariviere, V., & Milard, B. (2013). Cities and the geographical deconcentration of scientific activity: A multilevel analysis of publications (1987–2007). *Urban Studies*, 51, 2219–2234. doi:10.1177/0042098013506047
- Hausmann, R., & Hidalgo, C. A. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26), 10570–10575. doi:10.1073/pnas.0900943106
- Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16, 309–342. doi:10.1007/s10887-011-9071-4
- Hausmann, R., & Rodrik, D. (2003). Economic development as 'self discovery'. *Journal of Development Economics*, 72(2), 603–633. doi:10.1016/s0304-3878(03)00124-x
- Heimeriks, G., & Balland, P. A. (2016). How smart is specialisation? An analysis of specialisation patterns in knowledge production. *Science and Public Policy*, 43(4), 562–574. doi:10.1093/scipol/scv061
- Heimeriks, G., & Boschma, R. (2014). The path- and place-dependent nature of scientific knowledge production in biotech 1986–2008. *Journal of Economic Growth*, 14, 339–364. doi:10.1093/jeg/lbs052
- Heimeriks, G., & Vasileiadou, E. (2008). Changes or transition? Analysing the use of ICTs in the sciences. *Social Science Information*, 47, 5–29. doi:10.1177/0539018407085747
- Hidalgo, C. A., Klinger, B., Barabasi, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, 317(80), 482–487. doi:10.1126/science.1144581
- Kauffman, S. (1993). *Origins of order: Self-organization and selection in evolution*. Oxford: Oxford University Press.
- Kogler, D. F. (2015). Editorial: Evolutionary economic geography – theoretical and empirical progress. *Regional Studies*, 49(5), 705–711. doi:10.1080/00343404.2015.1033178
- Kogler, D. F., Heimeriks, G., Leydesdorff, L., Kogler, D. F., & Heimeriks, G. (2018). Patent portfolio analysis of cities: Statistics and maps of technological inventiveness technological inventiveness. *European Planning Studies*, 26, 2256–2278. doi:10.1080/09654313.2018.1530147
- Kogler, D. F., Rigby, D. L., & Tucker, I. (2013). Mapping knowledge space and technological relatedness in US cities. *European Planning Studies*, 21, 1374–1391. doi:10.1080/09654313.2012.755832

- Leydesdorff, L., Heimeriks, G., & Rotolo, D. (2016). Journal portfolio analysis for countries, cities, and organizations: Maps and comparisons. *Journal of the Association for Information Science and Technology*, 67, 741–748. doi:10.1002/asi.23551
- Maisonobe, M., Grossetti, M., Milard, B., Jégou, L., & Eckert, D. (2017). The global geography of scientific visibility: A deconcentration process (1999–2011). *Scientometrics*, 113, 479–493. doi:10.1007/s11192-017-2463-2
- Malerba, F., & Orsenigo, L. (2002). Innovation and market structure in the dynamics of the pharmaceutical industry and biotechnology: Towards a history-friendly model. *Industrial and Corporate Change*, 11, 667–703. doi:10.1093/icc/11.4.667
- Martin, R., & Sunley, P. (2007). Complexity thinking and evolutionary economic geography. *Journal of Economic Geography*, 7, 573–601. doi:10.1093/jeg/lbm019
- McCann, P., & Ortega-Argilés, R. (2013). Smart specialization, regional growth and applications to European Union cohesion policy. *Regional Studies*, 1–12. doi:10.1080/00343404.2013.799769
- Nomaler, Ö, Frenken, K., & Heimeriks, G. (2014). On scaling of scientific knowledge production in U.S. metropolitan areas. *PLoS One*, 9, e110805. doi:10.1371/journal.pone.0110805
- Rigby, D. L. (2013). Technological relatedness and knowledge space: Entry and exit of US cities from patent classes. *Regional Studies*, 1–16. doi:10.1080/00343404.2013.854878
- Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth*, 9, 131–165. doi:10.1023/B:JOEG.0000031425.72248.85
- Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). *Cambridge Journal of Regions, Economy and Society*, 11(1), 189–209. doi:10.1093/cjres/rsx024
- Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic Perspectives*, 8, 3–22. doi:10.1257/jep.8.1.3
- Stiglitz, J. (1999, January 27). Public policy for a knowledge economy. Retrieved August 7, 2019 from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.123.9173&rep=rep1&type=pdf>
- Storper, M. (1997). *The regional world: Territorial development in a global economy*. Guilford: Guilford Press.
- Tödtling, F., & Trippl, M. (2005). One size fits all? *Research Policy*, 34, 1203–1219. doi:10.1016/j.respol.2005.01.018
- UNESCO. (2010). *UNESCO science report 2010*. Paris: Author.
- Veugelers, R. (2016). *The European Union's growing innovation divide*. Bruegel Policy Contribution, No. 2016/08, Bruegel, Brussels.
- Wallace, M. L., & Rafols, I. (2015). Research portfolio analysis in science policy: Moving from financial returns to societal benefits. *Minerva*, 53, 89–115. doi:10.1007/s11024-015-9271-8
- Waltman, L., & Van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the Association for Information Science and Technology*, 63, 2378–2392. doi:10.1002/asi.22748