EPO-ARP Final Document

The interplay between science and invention networks in knowledge cohesion: Evidence from European Regions

İbrahim Semih Akçomak, Middle East Technical University
Umut Yılmaz Çetinkaya, TOBB University of Economics and Technology.
Erkan Erdil, Middle East Technical University
Müge Özman Gossart, Institut Mines-Télécom Business School

Abstract

We investigate knowledge convergence and knowledge cohesion in the European Union (EU) by analysing Framework Programme (FP) project collaborations and patent inventor collaborations from 2011 to 2019 at the NUTS2 regional level. Assuming that collaborations induce knowledge accumulation we differentiate between convergence and cohesion and conceptualize knowledge cohesion. Our empirical strategy is based on Social Network Analysis (SNA) and Simulation Investigation for Empirical Network Analysis (SIENA) where we analyse whether network structure affects regional characteristics as well as what factors and indicators affect network structure. The SNA and descriptive analysis show evidence for knowledge convergence indicating that knowledge-poor regions tend to catch up with knowledge-rich regions. SIENA results support evidence for convergence. We find that regional characteristics affect network structure but regions tend to collaborate with regions that are similar to themselves in terms of innovation level and general trust. The findings also indicate that network structure does not affect regional characteristics which reveals that collaboration led to a certain degree of knowledge convergence but not cohesion.

Keywords: collaboration, framework projects, patents, cohesion, SIENA

1. Introduction

This paper investigates knowledge convergence and knowledge cohesion in the European Union (EU) by analysing Framework Programme (FP) project collaborations and patent inventor collaborations seperately. In doing so, we attempt to merge two related, but somehow distinct, literatures on EU cohesion policy and collaboration-induced knowledge diffusion.

Investigating knowledge flows through patterns of collaboration between regions of the European Union (EU) is vital for building economic and social strategies, given the importance of "cooperation" in the EU agenda. EU Cohesion Policy is one such strategy that aims to strengthen economic, social and territorial cohesion by reducing disparities between EU regions. While there is a well developed literature on EU cohesion policy (e.g., Farole, Rodriguez-Pose and Storper, 2011; Bouayad-Agha, Turpin and Vedrine, 2013; Crescenzi and Giua, 2016; Di Cataldo, Monastiriotis and Rodriguez Pose, 2021) there is little research linking collaboration to cohesion (Hoekman et al., 2013; Lahdelma and Laakso, 2016; Boumans and Ferry, 2019). Moreover, there are various problems in measuring (especially social) cohesion (e.g., Jenson, 2010) and the empirical assessment whether the EU's cohesion policy achieved its economic targets (e.g., Dall'Erba and Fang, 2017; Ehrlich and Overman, 2020; Di Caro and Fratesi, 2021). This paper contributes to the EU cohesion literature in several aspects. First, in addition to economic, social and territorial cohesion we conceptualize knowledge cohesion and discuss how it differs from knowledge convergence. Second, this paper contributes to the empirical assessment debate by introducing Stochastic Actor Oriented Models (SOAM) and providing an application over FP project and patent inventor collaborations using Simulation Investigation for Empirical Network Analysis (SIENA). While our analysis helps to understand whether there is knowledge cohesion in EU regions it could be adopted for analysing economic and social cohesion as well. Third, by analysing collaborations to understand cohesion we extend the collaboration-cohesion link as stated above.

A crucial aspect of enhancing knowledge diffusion is inducing collaboration between EU regions. This strategy is both compatible with the development of the European Research Area (ERA) and the "strengthening research, technological development and innovation" thematic objective of the EU cohesion policy. There is a developing literature on the impact of collaboration-induced knowledge diffusion on outcome variables (e.g., Bergman and Maier, 2009; Wanzenbock, Scherngell and Brenner, 2014; Hazir, LeSage and Autant-Bernard, 2016; De Noni, Orsi and Belussi, 2018; van der Wouden and Rigby, 2019). For instance, Balland, Boschma and Ravet (2019) show that peripheral countries have become more integrated to the core by analysing FP project data from FP6 to H2020. This paper contributes to this literature in two aspects. First, to our knowledge this is the first research that uses SIENA to analyse collaboration networks. In this way we can analyse both the impact of regional characteristics on network structure and how network structure affects regional characteristics. Second, we compare two distinct collaboration networks: (i) FP project collaboration which we view as a science network, (ii) patent inventor collaboration network which we view as an invention network. Comparing these two distinct networks using a similar method brings robustness.

The findings of the social network analysis (SNA) indicates that there is an overall tendency for increased collaboration in the FP project network. Falling distance between regions in terms of project collaboration, increased clustering and network closure can be taken as signs of knowledge convergence. However, in the patent collaboration network there is a tendency for collaborations to weaken after 2015. The findings of the descriptive analysis reveals that there is persistence at the top-5 percentile regions meaning that knowledge hubs in 2011 remain to be knowledge hubs in 2019. Despite this finding, there is also an average trend of catching up, assuming that through collaborations regions accumulate knowledge. We find that in both the FP project network and the patent inventor network regions that are less endowed with knowledge tend to catch up with regions that are better endowed with knowledge, which can be taken as a sign of knowledge convergence. Finally, SIENA results show that while there are signs of knowledge convergence there is no evidence for knowledge cohesion.

This paper is organized as follows. Section 2 conceptualizes knowledge cohesion focussing on the difference between cohesion and convergence. Section 3 introduces data used in this research and the details of our empirical approach. The next section discusses findings under three subtitles, the findings of the social network analysis, the descriptive analysis and finally SIENA. Section 5 concludes by summarizing the findings and discussing policy implications in detail.

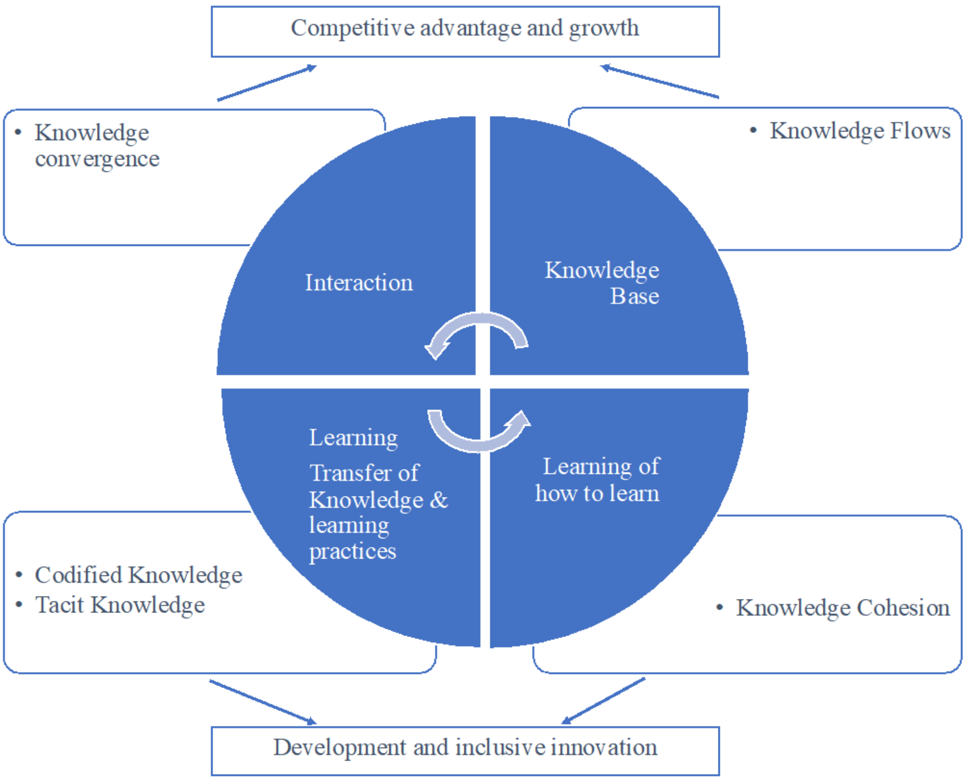
2. Knowledge convergence and knowledge cohesion

With the rise of evolutionary economics in the 1980s and 1990s, we observe a systematic and collective action-oriented account of treating knowledge in the literature (Nelson and Winter, 1982; Freeman et al., 1982; Freeman, 1987; and Lundvall, 1992). In addition to the individual dimension of learning, evolutionary economics has highlighted the processes of knowledge spillovers in the emergence of a collective knowledge base. The interaction through the collective knowledge base is assumed to enhance learning and learning of "how to learn". The systematic repetition of these interactions causes the genesis of knowledge networks within which not only codified but also tacit knowledge are transferred. Besides facilitating the transfer of knowledge among the actors, networks are also the means through which learning practices can be shared and transmitted. In other words, actors acquire the knowledge of "how to learn". This discussion highlights the difference between knowledge convergence and knowledge cohesion.

We take knowledge convergence to be mainly the outcome of three processes: 1) collaborations between knowledge actors, 2) Their mobility and 3) knowledge spillovers resulting from codified knowledge products like patents and academic publications. On the other hand, we take knowledge cohesion as the result of these processes accompanied by shared practices, institutions, and coherence in the innovation environment that knowledge actors are embedded in. In this sense, while both tacit and codified knowledge can be transferred in both contexts, the difference lies mainly in the extent to which the knowledge actors share a common institutional, social and economic framework. As a result, we contend that in a cohesive knowledge environment, it is not only knowledge that is exchanged per se, but also, and on a deeper level, the practices that give rise to the creation and evolution of knowledge.

Although it is generally assumed that collaboration leads to knowledge convergence, the extent of knowledge convergence has not been operationalised at broader levels and for different types of networks, such as project or inventor networks. In this study, we conceptualise knowledge convergence as an increase in collective knowledge and measure it as collaborative learning at the inter-regional unit of analysis. In addition to knowledge convergence this study provides evidence for knowledge cohesion and its socio-economic implications. Figure 1 summarizes this theoretical framework.

Figure 1: Theoretical Framework



2.1. Knowledge Convergence

Knowledge convergence is the growth of commonly shared knowledge that is brought by all the collaborating partners. It is the process in which two or more people share mutual understanding through social interaction, and the process of knowledge generation is believed to reflect its social nature (Nonaka, 1994; Brown and Campione, 1996; Hutchins, 1991, 1995; Ickes and Gonzales, 1996; Lave and Wenger, 1991; Resnick, 1991; Rogoff, 1998; Roschelle, 1992; Vygotsky, 1978; Webb and Palinscar, 1996; Teasley et al., 2008). Knowledge convergence is one of the most fundamental aspects of cognitive interdependencies among collaborators (Ickes and Gonzales, 1996; Roschelle, 1992;

Fischer and Mandl 2005). In this context, the dynamic capabilities of knowledge actors, referring to the skills, procedures, organizational structures, and decision rules are the key factors for generating value (Teece, et al., 1997). Convergence occurs because the mutual nature of collaboration leads to an increasing similarity in the cognitive representations of group members (Jeong and Chi, 2007). The literature on knowledge convergence is generally based on the perceptive definition of knowledge convergence and present qualitative studies for the understanding of mechanisms behind it (e.g., Dretske, 1981; Azmitia, 1988; Forman and Cazden, 1985; Brown and Duguid, 1991; Nonaka; 1994; Hutchins, 1995; Graesser, et al., 1995; Roschelle, 1992; Roschelle and Teasley, 1995). We contend that to understand convergence we can benefit from collaboration induced networks.

Based on this notion of common knowledge, we define knowledge convergence to be an increase in common knowledge following collaboration. Moreover, reminding the three processes that lead to convergence above, to the extent that collaboration is responsible for knowledge convergence, the level of convergence would depend on interactions between knowledge actors. Thus, the process is endogenous with a bidirectional causality and continuous feedback loops reactivate new opportunities for the agents in the process. At the micro level, one of the effects of increased collaborations between organizations is their convergence in the knowledge space, rendering them more similar to each other in terms of common technological competences (Mowery et al., 1998). However, there is also diminishing returns from such interactions, because as convergence increases learning that takes place between organizations reduces as they can add less and less to each others' knowledge (Mowery et al., 1998; Gilsing et al., 2008). Accordingly, one of the findings in the literature is that an inverted-U relationship exists between technological distance between actors and their learning (Mowery, et al., 1998; Gilsing et al., 2008, Schoenmakers and Duysters, 2006; Nooteboom et al., 2007). However, this strand of research often measures knowledge through bibliometric analysis. Thus, the aspects of knowledge cohesion between actors are left unexplored.

The new opportunities created by the interactive learning spaces can be described as the successful innovation activities which utilizes the common knowledge silos created by the above-mentioned interactive process. However, the mechanism is not automatic and/or self-adjusting and there needs to be a menu of framework conditions and specific characters of cognitive processes based in regional/local ecosystems. The capability of regional ecosystems to integrate to an upper-level system such as national, supranational and global systems determine sustainability of success of these systems. Moreover, the dynamic nature of this articulation determines the co-evolution of the whole system and establishes hot spots as evident from a recent study on global networks (WIPO, 2019)

2.2. Knowledge Cohesion

Knowledge has always been an endogenous element of cohesion, and both knowledge convergence and cohesion are strongly linked with each other. Just as there could be knowledge cohesion without convergence, the two processes can also accompany each other. For example, some segments of regional economies can be highly cohesive, sharing similar contexts, cultures and capabilities, yet draw upon completely different knowledge bases, as in India, Turkey and Mexico. However, convergence does not necessarily guarantee cohesion. Knowledge convergence emancipates the flows of knowledge through the inter- and intra-network relations whereas it has the potential to create new traps of enslavement if the framework conditions for cohesion do not historically exist or are created to some extent. Therefore, knowledge cohesion is not simply deduced to a linear process. It includes the non-linear and complex relations of transferring learning practices and mechanisms of how to learn. For example, without the existence of sales and marketing knowledge, the scientific and technological knowledge may be useless. Another possibly more chaotic and complex social process is the transfer of knowledge cohesion to social and economic cohesion through which several complicated mechanisms among the acts and actors are observed. The initial definition of knowledge cohesion is provided as the unification of combinatorial knowledge bases as a whole which makes the knowledge convergence as a sufficient condition of the process towards knowledge cohesion.

Based on the above discussion, it is possible to summarize knowledge convergence and cohesion on a two-dimensional taxonomy as presented in Figure 2. It is important to note that we focus on the macro level, which refers to cohesion between different regions in an innovation system. In doing so, it will be useful to clarify how we distinguish between knowledge convergence and knowledge cohesion on a two-dimensional axis.

We take knowledge convergence in terms of the common knowledge bases. Regions with low knowledge convergence (low common knowledge) correspond to the cases where significant divergence is expected in terms of the intensity and level of knowledge diffusion and generation activities. For example, regions' participation in global knowledge networks, the levels of public and private R&D activities, innovative outputs etc., may exhibit significant differences, marked distinctly as a gap between high and low performing regions. On the other hand, regions with high knowledge convergence will be more similar to each other, in terms of the indicators of overall science and technology performance.

When it comes to knowledge cohesion, on the other hand, our focus shifts from the output indicators, to a process view, whereby we focus on the general institutional space that shapes science, technology and innovation activities. In this sense, when different regions are highly cohesive, they would have shared norms, values and institutions in their systems of science and technology, as compared to other regions, whether they are connected through network ties or not (in most cases there will be increased networking activities as well). In a way, cohesiveness implies that innovative actors are bound together by tangible or intangible ties, and their communication is largely facilitated through shared norms and values in knowledge practices, in the generation and diffusion of knowledge. By engendering trust, such cohesiveness largely facilitates processes of knowledge diffusion and sharing.

Figure 2. Taxonomy of knowledge dynamics: Convergence and cohesion

Convergence / cohesion	**High**	**Low**
High	**KNOWLEDGE SPACE II** High convergence, high cohesion **Network characteristics:** dense networks among similar actors **Risks:** Difficulty of novelty creation, over-embeddedness, lock-in, information overload **Benefits**: Reduced transaction costs, low risk of wrongly performing an R&D and innovation activity	**KNOWLEDGE SPACE III** High convergence, low cohesion **Network characteristics:** largely isolated and possible competitive clusters with (often) weak ties between them **Risks:** lack of common broad institutional frameworks prevent smooth knowledge flows, trust, social capital, etc. **Benefits**: generation of novelty due to coexistence of diversity on one hand, and common (field specific) knowledge base on the other. Weak ties between clusters can promote novelty generation
Low	**KNOWLEDGE SPACE I** Low convergence, high cohesion **Network characteristics:** core - periphery **Risks:** increased inequality and access opportunities to S&T **Benefits**: better potential for knowledge flows due to shared practices and institutions, and better potential creation of novelty due to new ideas that can arrive from the periphery. Sharing of knowledge is better supported.	**KNOWLEDGE SPACE IV** Low convergence, low cohesion **Network characteristics:** isolation **Risks:** sunk costs in R&D and innovation activities, not able to follow the common path of technological development **Benefits:** Learning by doing processes may help indigenous innovations

In conclusion, while there is a rich literature on knowledge convergence, and, it is surprising that the concept of knowledge cohesion has not been developed before. As for knowledge convergence, it refers to the extent to which various actors form, access and benefit from a common pool of knowledge stock. At the same time, cohesion, as defined in different contexts, often highlights the forces that keep actors together in a system. Conceptualizing knowledge cohesion is important, because it is essentially a range of practices in knowledge generation and diffusion that lie behind the success of many regions, and explored within a scattered array of disciplines. These practices cannot be well understood with reference to knowledge convergence alone, as was often the case in the literature.

There are several premises on which our discussion is built upon. First, we take knowledge dynamics as the main force behind change in societies and economies. Second, collaborations between different actors in an innovation system are the main motor behind knowledge flows. Third, our focus is not only on the mere transfer of knowledge, but also on the specific processes through which actors learn how to learn. Fourth, we contend that a cohesive knowledge environment implies, over and beyond convergence, the processes through which actors share common norms and values, create environments of trust and open knowledge sharing, and are able to maintain both homogeneity and diversity in the knowledge systems. As a result, an exploration of the extent to which a knowledge environment is cohesive requires thinking of cohesion along with several other factors that have been deemed as critical in the literature. These are, a reconceptualization of the role of proximity, of social capital and networks, of uncertainty and the way it is coped with in the innovation systems, as well as a better understanding of how actors are equipped with capabilities to create new paths towards sustainability in the face of path dependent trajectories. The empirical application of this paper will consider the functioning of the concepts outlined in this section as the existing data paves the way for such an analysis. The long time series data on FPs and patents certainly provides evidence for testing the existence of knowledge convergence while it may also present some considerable evidence for knowledge cohesion.

a. 2.3. The role of research networks in knowledge convergence and cohesion

In practical terms, in this paper we take the policies aimed at reducing the knowledge gap between star regions (i.e., the knowledge hubs) and others to refer to knowledge convergence. Therefore, in line with the upper part of Figure 1, increase in common knowledge and similarity between S&T indicators would imply convergence. On the other hand, given the importance of the transfer of learning practices (the bottom part in Figure 1), we contend that a distinction between convergence and cohesion is of paramount importance.

In investigating these issues, the evolution of networks can provide important insights. First, convergence will be revealed by a region's integration into knowledge networks over time. For example, initially peripheral regions integration into the core of the network would be an indicator that the gap between regions are reducing, thus signalling increased convergence. However, such an analysis may not be sufficient to analyse the extent to which regions become more cohesive in terms of the convergence in their scientific and knowledge based institutions, and their capabilities to develop practices about second order learning (learning how to learn). For example, some regions collaborate more than others, and thus they converge in terms of common knowledge, with their partners. However, this does not mean that the region's cohesiveness with others will be high. Some regions, although they converge, can have a big gap with partners in terms of their capacity and capabilty that enable transfer of skills and capacities (second order learning). So as to have a better understanding of these, we argue that the drivers of network evolution can reveal some insights.

The transfer of skills and capabilities is important especially given that many EU regions are unable to use funds effectively to boost innovation, as at one point institutions become a constraining factor in their ability to use the funds (Schmidt, 2019). Skills and funds can be better transferred with similar institutions. When institutions are very different in terms of capabilities, new skills acquired will be difficult to assimilate. A region which does not have capabilities and relevant institutions to build upon what it learns, might lag behind in terms of long term innovation performance, although it participates in research networks. For these purposes, in the next sections, we analyse the drivers of networks in science and invention networks. We compare structural characteristics and analyse the differences in the drivers of their evolution so as to develop some insights about convergence and cohesion dynamics.

3. Methodology and data

The empirical strategy rests on three steps: (i) using social network analysis (SNA) to analyse collaboration networks over time, (ii) employing centrality indicators from the SNA to provide descriptive statistics and simple Ordinary Least Squares (OLS) estimations to assess knowledge convergence, (iii) using Simulation Investigation for Empirical Network Analysis (SIENA) to investigate the impact of regional characteristics on network structure and the impact of network structure on regional characteristics simultaneously to assess knowledge convergece and cohesion. Before discussing each step further, section 3.1 presents information on data and how networks are formed.

3.1. Data

For the analysis, two separate data sources were used: (i) CORDIS database, which contains the projects supported by the European Commission (EC) which we used to form the framework project (FP) network, (ii) the patent data provided by PATSTAT which we used to form the patent network. Both databases were firstly cleaned, separated by years (2011, 2013, 2015, 2017 and 2019) so that they match the Regional Innovation Scoreboard data and then converted into network data.

In CORDIS[1], nodes of the FP network were obtained by using the address information of each project partner (*i.e.*, the coordinator and all partners in a project) and aggregating to the NUTS2 regional definitions.[2] Projects acted as the links established among nodes. For instance, if a project has 3 partners, then a total of 3 links were established among these partners. Links were set up with no direction because we assume that knowledge transfer will be mutual. It is extremely difficult to know the direction of knowledge flow even if there are large observable differences in terms of knowledge stock among the geographies of partners. Patent network was also set up in a similar fashion to the FP network; in this case, the links between nodes were formed by the partnerships of inventors.

As stated above we specifically limit our analysis to the period 2011-2019 because RIS data at the NUTS2 level is only available in those years.[3] RIS contains indicators as well as a regional innovation index starting from 2011 for every two years. The regional innovation index is used as a proxy for the level of innovation in a region.

Two types of analysis, a standard Social Network Analysis (SNA) and Simulation Investigation for Empirical Network Analysis (SIENA) were carried out using the data sets above. While in SNA, only CORDIS and PATSTAT databases were used, in SIENA we benefited from all the databases listed in Table 1. By definition a project partnership continues as long as the project continues, whereas collaboration in a patent is observed only when the patent is granted. To be specific, in the project network, the link is sustained until the end of the project. For instance, if a project started in 2011 and ended in 2015, the collaboration relationship among the partners in this project is kept in 2013 and 2015 as well as 2011. On the other hand, if two inventors collaborated in a patent granted in 2011 then the collaboration relationship between the inventors is only kept in 2011 unless the same inventors are granted another patent. Thus the patent network is expected to be less dense compared to the FP project network not only because collaboration in patents is rare compared to collaboration in projects but also because of how the network is formed.

Table 1: Databases used in the network analyses

Databases	**CORDIS**	**PATSTAT**
	Table 5	Table 6
RIS, EUROSTAT, ESS	Table 9	Table 10

As stated earlier our empirical strategy can be summarized in three steps. In the first step, CORDIS and PATSTAT were used to form separate project and patent networks. In the

[1] https://cordis.europa.eu
[2] Regional statistics by NUTS classification (reg), https://ec.europa.eu/eurostat/web/regions/data/database
[3] https://ec.europa.eu/info/research-and-innovation/statistics/performance-indicators/regional-innovation-scoreboard_en

network each node is a NUTS2 region and links between nodes represent either collaboration among project partners (FP network) or collaboration among inventors in a patent (patent network). Section 4.1 presents the results of the simple SNA. Then we used two common centrality indicators (degree and betweenness), brought them into the same measurement unit and provided descriptive analyses comparing the two different networks so as to make a preliminary assessment of knowledge convergence. The results of these descriptive analyses are presented in section 4.2. Finally in section 4.3 we present the results of the SIENA separately for the FP project network and patent network. The list of the data and variables used in the analysis, the descriptions and their sources are provided in Table 2.

Table 2: The data, variables and sources

Data/variable	Source
Framework Project collaboration	Framework project data is freely available from CORDIS. https://cordis.europa.eu/projects/en. We used the partners of each project to determine collaboration among partners which then aggregated to regional level using address information of each partner.
Patent collaboration	PATSTAT. https://www.epo.org/searching-for-patents/business/patstat.html. We used names and address information of the inventors. If there are two or more names as inventors in a patent we assumed that there is collaboration between persons which then aggregated to regional level using address information of each inventor.
Regional innovation index (innovation)	The Regional Innovation Scoreboard provides a regional innovation index from 2011 onwards available every two years. https://ec.europa.eu/info/research-and-innovation/statistics/performance-indicators/regional-innovation-scoreboard_en
General trust (trust)	Question from European Social Survey (ESS) measuring general trust "Most people can be trusted or you cannot be careful". Answers range from 0 (you can't be too careful) to 10 (most people can be trusted). Data available in 2010 to 2018 every two years, with a total of five waves. https://www.europeansocialsurvey.org/data/download.html?r=5
Human resources in science and technology (employment)	Persons with tertiary education and/or employed in science and technology as a percentage of population available from EUROSTAT. https://ec.europa.eu/eurostat/databrowser/view/HRST_ST_RCAT__custom_1665276/default/table?lang=en
Patent applications	Patent applications per million population. 2010 data available from EUROSTAT. https://ec.europa.eu/eurostat/databrowser/view/pat_ep_rtot/default/table?lang=en
Log Population	Logarithm of population 1st of January. 2011 data available from eurostat. https://ec.europa.eu/eurostat/databrowser/view/demo_r_d2jan/default/table?lang=en

3.2. Empirical strategy

The first step in our empirical strategy is based on analysing Framework Programme (FP) project collaborations and patent collaborations using SNA. How FP project collaboration and patent collaboration networks are formed are already explained above. In SNA a new network is formed every period. Thus we have five separate networks for 2011, 2013, 2015, 2017 and 2019 based on either FP project or patent inventor collaborations. Using SNA statistics we compare the project network to the inventor network.

In SNA, the node properties/characteristics are not taken into account and the analysis is conducted based only on the existence or absence of links. In other words, the factors that could have an affect on the establishment and termination of these links are not considered. Such an analysis looks at whether links exist at a given time t. For instance, a region's links to other regions in the form of project collaborations can change over time. But at the same time regional characteristics, such as innovation and employment may also change. SNA does not consider the impact of network structure on regional characteristics or vice versa. SNA also does not analyse structural changes in the network (i.e., how the network structure has changed from 2011 to 2013 and how the 2011 network affected the 2013 network and so on). Section 4.1 presents the results of SNA for project collaborations and inventor collaborations separately. This analysis mostly gives clues about knowledge convergence.

The second step of the empirical strategy is based on obtaining two commonly used centrality indicators (degree and betweenness centrality) from the SNA for each year for each collaboration network and bringing them into the same measurement unit such that they are comparable over the years. Because the network of a particular year, say 2011, is different from another year, say 2015, the network statistics in a period cannot be compared to another. To circumvent this problem we used percentile ranks such that each "centrality indicator-year-collaboration network" combination is associated with a percentile rank. For instance, degree centrality values for 2011 are available for 280 NUTS regions in the FP project network. Each of these 280 regions has a percentile rank that ranges from 1 to 100, higher values indicating better position in the network regarding the number of links a region has. When this process is replicated to other years we can compare the percentile rank of a region for 2011, 2013, 2015, 2017 and 2019. Thus we roughly have an idea regarding the position of the region over time. This information is utilized in several ways.

First, we can list top regions according to different centrality indicators over time. Looking at top regions enables us to see whether there is persistence in knowledge creation assuming that better position in a collaboration network is associated with higher knowledge accumulation. It could be the case that starting levels are important such that knowledge hubs in 2011 still continue to be knowledge hubs or that through collaborations some regions accumulate knowledge and proceed to become a knowledge hub. The latter is what we expect to find, the former is a finding against knowledge cohesion as it basically states that starting levels are important. But this analysis only looks at the top regions. Second, we associate the difference in percentile ranks from 2011 to 2019 to the starting level percentile rank. We want to see the simple correlation between the starting levels and changes over a period. A negative correlation coefficient may be a hint toward knowledge convergence. Third, inspired by the simple empirical economic growth model we regress changes in percentile rank on starting level percentile rank, log population to control for size, patent applications per million inhabitants to control for initial knowledge stock and country dummies to control for country fixed effects. We expect to find a negative coefficient for the starting level percentile rank to talk about knowledge convergence. Such a finding would tell us that regardless of size, initial knowledge stock and country effects, an average region obtains a better position in the collaboration network over time. While this finding tells nothing about the changes in the network structure and what regional factors affect this change or whether network structure affects regional characteristics, it tells us that some regions are able to take advantage from collaboration to better link with the network which may result in knowledge convergence.

The third step involves estimation of a Stochastic Actor based Model (SOAM) using SIENA which considers both the links (changes in the network structure) and the properties of the nodes (regional characteristics). SIENA consists of two basic components: networks (in our research, structural changes of the network over time) and attributes (in our research, regional characteristics such as the innovation level of the region).

SIENA is a statistical tool, which was developed to analyse longitudinal network data such that networks are observed in different periods. (Snijders; 1996; 2001; Snijders, van de Bunt and Steglich, 2010). It is "a set of methods implemented in a computer program that carries out the statistical estimation of models for repeated measures of social networks according to the Stochastic Actor-oriented Model (SAOM)" (Ripley et al., 2021). Similar to other SAOMs, SIENA allows quantifying the evolution of the network between different time periods. SIENA obtains the change in the links between actors using a network dynamics approach. In other words, just as in the standard SNA, links are established, maintained, and broken in SIENA as well. But when making the calculations, SIENA takes into consideration both the network structure (i.e., how the collaboration network changes) and the characteristics of the nodes (whether regional characteristics affect network structure or vice versa).

SIENA has already been used in various research especially in the psychology literature to investigate friendship ties, smoking behaviour, alcohol use and bullying (Cheadle and Goosby, 2012; Wang et al., 2015; Leszczensky and Pink, 2015; Shin, 2017; Gremmen et al., 2017; Hooijsma et al., 2020). Kalish (2020) provides a decent first introduction of SAOM using SIENA to management scholars who especially work on organizational research. In this research we used SIENA to analyse co-evolution of one-mode networks (collaborations in FP projects and collaboration in patents separately) and individual behaviour (characteristics of the regions). This can be viewed as longitudinal network data where one or more changing nodal variables can also be treated as dependent variables, referred to as behaviour. In such a setting the network will influence the dynamics of the behaviour and the behaviour will influence the dynamics of the network.

The network data that is used in step two above was reassembled so that SIENA estimations could be performed in the R environment. The network data comes from FP project and patent inventor collaborations. For the attributes we used three indicators: regional innovation index as a proxy for innovation, generalized trust as a proxy for social capital and human resources in science and technology to control for size.

When performing analyses in SIENA, a closed network and at least two time periods are required. It is possible to conduct two types of analyses in SIENA. First, one can perform a structural analysis similar to the SNA which gives information about the density, reciprocity and transitivity regardless of node characteristics (i.e., over time how the network structure changes?). SIENA's distinctiveness comes from the fact that node characteristics can be included into the analysis. Node characteristics can be defined as attributes and/or behaviour in the system. It could be the case that the node characteristics affect the nodes' establishing, sustaining, and terminating links in the network. If this is the case, node characteristics are defined as attributes. For instance, to find out whether innovation level of a region affects a region's propensity for establishing links in the network, we include the innovation level of each region as an attribute. In our analysis we included innovation level, general trust and human resources in S&T of a region as attributes. In this way, in addition to network effects such as density, transitivity etc. we can also analyse how regional characteristics affect network structure. It could also be the case that the network structure affects node characteristics. For instance, we may want to know whether a region emulates or tries to reach to an innovation level as those regions they establish link with (behaviour). If this is the case then we can assume that the regional characteristics are changing in response to changes in network structure which we take as a sign for knowledge cohesion.

Table 3 presents the effects included in SIENA in this research, the graphical displays and the interpretation of each effect. In this research we take innovation level and general trust as both attribute and behaviour variables, human resources in science and technology as an attribute only to control for size, and distance[4] among nodes as an attribute only to control for geography.

Finally, Table 4 presents the research questions of the EPO-ARP project proposal and associates each research question to a network analysis in this research. The question regarding network dynamics is addressed by both the SNA and SIENA. The impact of regional characteristics on network structure is mainly analysed by using SIENA. Addition to the questions above, we also investigate the impact of network structure on regional characteristics to comment on knowledge cohesion as well as knowledge convergence.

[4] Calculated using Territorial Typologies (TERCEF) https://gisco-services.ec.europa.eu/tercet/flat-files.

Table 3: Effects included in SIENA and their interpretation

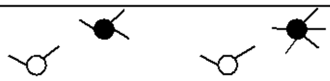
Transitions in the network (t → t+1)	Effect	Interpretation
	Network Dynamics: the impact of regional characteristics on the network structure	
	Density	tendency to form arbitrary (if coeff. is positive) or selective (if coefficient is negative) collaboration
	Transitivity	tendency for region's collaboration partners to collaborate
	# of links (node)	the more collaborations a region has the more it attracts
	Strong tie	previous collaboration with a region drives further collaboration
	X ego	region that has higher value in X tend to form more collaborations
	X similarity	collaborations occur more often between regions with similar value of variable X
	Behaviour Dynamics: the impact of network structure on the regional characteristics	
	linear shape	tendency for region to engage in behaviour
	quadratic shape	tendency for changes in behaviour to depend on initial levels
	X degree	Effect of own activitiy in the network on behaviour
	X average alter	tendency for regions to have values of variable X similar to those whom they have collaborated. Effect of neighbours average behaviour (measures contagion)

Table 4: Research questions of the EPO-ARP project

Research Questions	SNA FP Project Table 2	SNA Patent Table 3	SIENA FP Project Table 6	SIENA Patent Table 7
Do patent and research networks have the same dynamics? If not, what can be the sources of differences?	X	X	X	X
Do patents have any impact on the establishment of research networks?				X
Do the innovation performances of regions affect the establishment of patent networks?				X
Do the innovation performances of regions affect the establishment of research networks?			X	

4. Results

The findings of our research are presented under three subtitles. First we present descriptive network statistics of each period (2011 to 2019, five waves) and compare science networks (FP project collaborations) to invention networks (patent collaborations). Looking at simple statistics like degree, betweenness, closeness and eigenvector centrality may show trends in access to and importance within the network and even bridging roles. Then by bringing different network statistics to the same measurement unit we look at the positional changes in a network over the years. This simple descriptive analysis shows top regions in terms of collaboration and may give clues about persistence and may even show emergence of new knowledge hubs. Moving from the idea of convergence to cohesion needs a panel network analysis where one can analyse the impact of regional characteristics on network structure and in turn how network structure affects the characteristics of the regions. Thus, finally we present the findings of SIENA for the science network and invention network seperately.

4.1. Network Statistics

As seen in Table 5, the FP project network consists of 281 nodes. Various network indicators reveal that collaborations between NUTS2 regions have increased over time. The distance among nodes (average geodesic distance) decreased between 2011-2019. When the entire network is taken into consideration, the distances between NUTS2 regions in terms of cooperation and information flow, have decreased. In addition, there is a slight increase in network density, which shows the total number of actual links in a network with respect to the total number of possible links, if all nodes were connected to each other. As in density, we observe an increase in the number of links per node (average degree) as well. Average betweenness centrality indicates the extent to which a node in the network acts as a bridge connecting other nodes, in a way showing the node's importance in the information flow between unconnected nodes. This value is highest in 2015 and then decreases. This situation shows that relatively less new links were established based on these bridging nodes, which demonstrates that the network structure started to become more distributed. Closeness centrality measures the average shortest distance from each node to the other node. As this value increases, nodes become closer to each other. It can be seen that this value is stable throughout the years. Eigenvector centrality shows the extent to which a node is connected to other important nodes in the network. It's value remained almost unchanged over the years in the FP project network. The last indicator is the clustering coefficient. Clustering coefficient focuses on the egocentric networks rather than the entire network, and it shows the extent to which alters of a node are connected among themselves. In other words, it is calculated as: [the number of links connecting a node's neighbors] / [the total number of possible links among the node's neighbors]. For instance, assume that A has three neighbors: B, C, and D. There are links between B and C as well as B and D. Clustering coefficient is 2/3 or 0.66. The clustering coefficient is calculated like overall network density, but only using a subset of nodes. The average value of the clustering coefficient, calculated separately for each node, tends to increase. This demonstrates that there is a tendency for network closure.

Table 5: Network statistics - FP project collaborations

	2011	2013	2015	2017	2019
Graph Type	Undirected	Undirected	Undirected	Undirected	Undirected
Vertices	281	281	281	281	281
Unique Edges	25287	25509	25297	25899	26616
Edges With Duplicates	0	0	0	0	0
Total Edges	25287	25509	25297	25899	26616
Self-Loops	255	254	255	260	258
Average Geodesic Distance	1.3543	1.3486	1.3587	1.3388	1.3205
Graph Density	0.6363	0.6420	0.6366	0.6517	0.6700
Average Degree	179.9786	181.5587	180.0498	184.3345	189.4377
Average Betweenness Centrality	49.9253	49.1352	50.8968	47.7616	45.2064
Average Closeness Centrality	0.0027	0.0027	0.0027	0.0027	0.0028
Average Eigenvector Centrality	0.0036	0.0036	0.0036	0.0036	0.0036
Average Clustering Coefficient	0.8389	0.8428	0.8409	0.8371	0.8413

The patent network also consists of 281 nodes as can be seen in Table 6. As there are anomalies in 2019 data, the analysis is performed for the years 2011-2017. In terms of the variables, there seems to be an increase in collaborations until 2015, and decrease afterwards. For example, the distance among nodes (average geodesic distance) falls between 2011-2015 and increases in 2017. The density of the network also tends to fall after 2015, demonstrating that the patent collaboration links among nodes decrease. Besides, a decrease is observed in the number of links per node (average degree) as well. Average betweenness centrality peaked in 2015 and then decreased, in line with the observation that the geodesic distances between nodes have increased. Closeness centrality, eigenvector centrality, and clustering values demonstrate almost no change over the years. In summary, this network shows that there is a decrease in patent collaborations between NUTS regions, especially after 2015. This pattern is different from the FP projects, in which we observed an overall tendency for increase in collaborations during the period 2011-2017.

What are the implications of these preliminary statistics for knowledge convergence and cohesion? Firstly, it is possible to observe that, although there are similarities between the two networks, there are also divergences, notably after 2015, in the evolution of the networks. In the case of FP, regions tend to converge to each other, as the network statistics reveal. For example, reduced geodesic distances, higher network density and average degrees of nodes, and higher clustering coefficients can be interpreted as signs of knowledge convergence. On the other hand, while we observe similar patterns in the patent network initially, after 2015, there are signs of loosening of networks. Stated differently, geodesic distances between regions increase, network density as well as average degrees fall. These signal that the two networks are possibly driven by different mechanisms of evolution in time. In section 4.3. below, we further explore these mechanisms (especially the discussion of Table 11).

Table 6: Network statistics - Patent collaborations

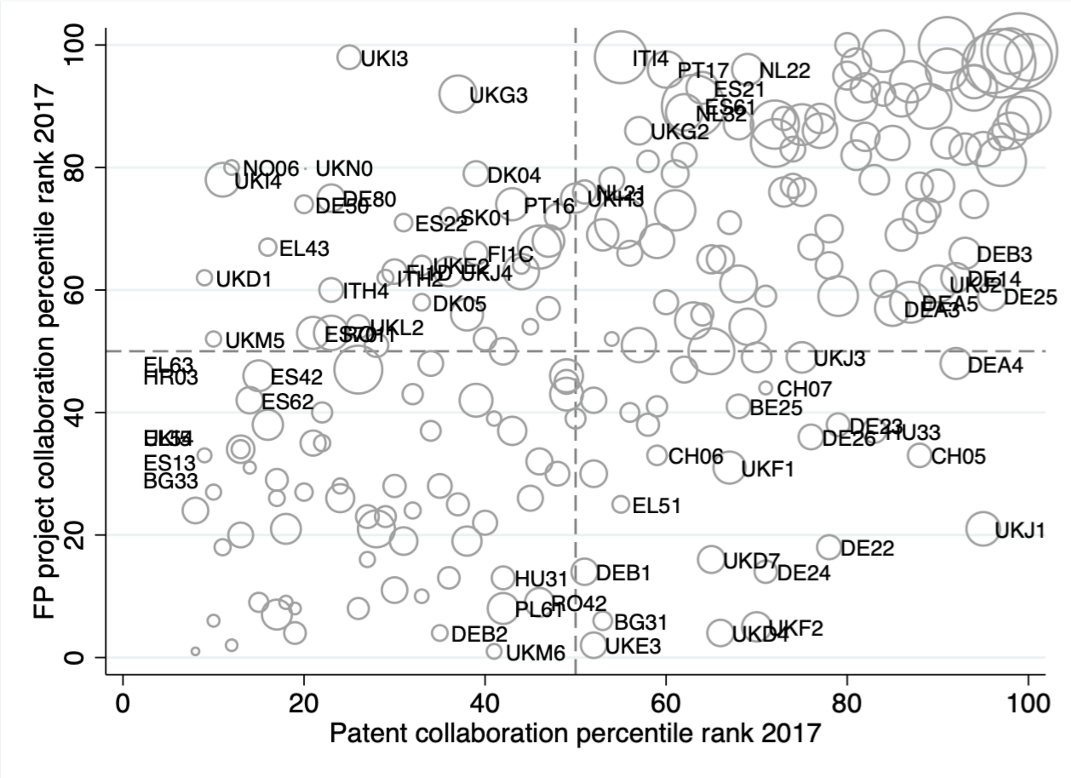
	2011	2013	2015	2017	2019
Graph Type	Undirected	Undirected	Undirected	Undirected	Undirected
Vertices	281	281	281	281	281
Unique Edges	3802	3996	3952	3635	1357
Edges With Duplicates	0	0	0	0	0
Total Edges	3802	3996	3952	3635	1357
Self-Loops	216	219	222	222	171
Average Geodesic Distance	2.0715	2.0621	2.0388	2.1059	2.8248
Graph Density	0.0912	0.0960	0.0948	0.0868	0.0301
Average Degree	27.0605	28.4413	28.1281	25.8719	9.6584
Average Betweenness Centrality	96.9253	96.9359	94.8185	101.8149	123.8505
Average Closeness Centrality	0.0018	0.0018	0.0018	0.0017	0.0013
Average Eigenvector Centrality	0.0036	0.0036	0.0036	0.0036	0.0036
Average Clustering Coefficient	0.3831	0.3904	0.3920	0.3850	0.2868

4.2. Descriptive results

Establishing and maintaining knowledge stock may present advantages to regions. On one hand a region may invest in infrastructure and human capital to start accumulating knowledge. This effort needs exploration activities either through transferring codified knowledge or through involvement in research and innovation networks. The initial knowledge stock may further be enhanced by exploitation as well as exploration activities. While the initial knowledge stock of a region creates an impetus for exploration and exploitation activities (i.e., extended network in terms of links to different regions), the position of a region within a network may further enhance its knowledge stock. Thus in terms of our descriptive analysis we expect a certain level of persistence for some regions that are bigger in size and engaged in the network at an earlier time.

But it could also be the case that some regions, through the mechanisms discussed above, are successful in creating and maintaining the knowledge stock even though they had certain initial disadvantages compared to other regions that for long have fueled research and innovation activities. Size could create a disadvantage, initial infrastructure and human capital or geography may create certain disadvantages. But still some peripheral regions (either within a country or throughout Europe) have developed in terms of knowledge stock compared to others. In our descriptive analysis we also look for emerging knowledge hubs as well as persistent knowledge hubs. Our assumption is that through collaboration activities by both participating in FPs and/or inventor collaboration in patents, regions build knowledge stock. However the position of a region within a collaboration network and the change of its position through time determines (the pace of) knowledge convergence.

Figure 3: Correlation of FP project collaboration and patent collaboration percentiles in 2017



In section 4.1 we already discussed descriptive network statistics and what they mean in terms of knowledge convergence and cohesion. However, to compare a region's position within a changing network structure we need to bring network statistics in a common measurement unit. For instance, Oberbayern (DE21) has different centrality statistics over the years; however by just looking at the numbers we cannot analyse Oberbayern's position in a network over the years. In each period (2011 to 2019) network statistics are computed over a different network where the number of nodes and the links between them are different. To bring centrality statistics to a common measurement unit we used each region's centrality statistics (degree, betweenness, closeness and eigenvector centrality) and computed percentile ranks for centrality statistics in all periods (2011 to 2019 every two years). Thus a percentile rank (say 95th) in degree centrality in 2011 shows the position of a region in the network of 2011 according to degree centrality. The rank of a region through time is comparable. For instance Oberbayern (DE21) is top 5 percentile region in both FP project collaborations and patent collaborations in all five periods from 2011 to 2019 which shows persistence. Even though the network structure is different in different periods, Oberbayern is always a top-5 percentile region (according to network degree). But for instance even Wien (AT13) is not in the top-5 percentile list according to degree centrality, its percentile rank ranges from 93 to 97 over the years for FP project collaborations and from 83 to 89 for patent collaborations also displaying persistence. Figure 3 shows the correlation between FP project collaboration and patent collaboration weighted by the size of the region. The correlation is 0.66 and is significant at the 1% level. Larger regions tend to be good both in terms of FP projects and patent collaborations. The northwest of the figure displays regions that are more successful in collaboration in science and the southwest of the figure displays regions that are good in participating in invention networks but are poor in participating in scientific networks.

Table 7: Top 5 percentile NUTS 2 regions (according to network degree) - FP project collaborations vs. patent collaborations

FP project collaboration					Patent collaboration				
2011	2013	2015	2017	2019	2011	2013	2015	2017	2019
ITI4	ITI4	ES30	ES30	FR10	DE21	DE21	FR10	DE21	DE21
FR10	FR10	FR10	BE10	ES30	DE71	FR10	DE21	DE11	DEA1
BE10	DE21	ITI4	FR10	BE10	DEA2	DE71	DE71	DEA2	DE71
ES30	ES30	BE10	ITC4	ITI4	FR10	DEA2	DEA1	DE71	DE11
EL30	UKI3	DE21	UKI3	NL33	DE12	DE12	DEA2	DE12	DE12
ITC4	ES51	UKI3	NL33	ITC4	DE11	DEA1	DE12	FR10	DEA2
UKI3	BE10	ES51	ES51	ES51	DEA1	DE11	CH04	DEA1	DE30
DE21	NL33	ITC4	ITI4	EL30	DEB3	CH01	CH03	DE14	DE25
ES51	ITC4	AT13	AT13	UKI3	CH04	DE30	DE11	CH03	FR10
NL33	EL30	EL30	EL30	ITH5	DE30	UKJ1	UKJ1	ITC4	CH04
DK01	SE11	NL33	PT17	PT17	ITC4	DEB3	DE13	DEB3	DE13
DEA2	NL22	DK01	FI1B	DE21	DE25	CH04	DEB3	CH04	DEA3
	FI1B	FI1B	DE21	AT13	DE13	DE13	UKH1		
	DEA2		DEA2	FI1B					
	UKG3			DEA2					
				ITI1					

Table 7 lists the top 5 percentile NUTS 2 regions according to the network degree where the left panel of the table lists regions according to collaboration in FP projects and the right panel lists regions according to patent collaboration. Network degree is the number of links of a region to other regions reflecting opportunity. Once those links are formed there is a tendency of collaboration to continue and even deepen if the relation between regions builds trust. Several observations can be made from Table 7. First, patent collaboration is more homogenous. Most top-5 percentile regions are from Germany. Regions from Switzerland, United Kingdom (UK), France and Italy are also present in different periods. However, the list of top 5 percentile regions in FP project collaborations is more heterogeneous. There are regions from 13 different countries. A second observation is that most regions that are listed top-5 in 2011 are all the time at the top-5 percentile list. 10 out of 12 regions that are considered to be top-5 region in FP project collaboration are in the list for all years. When we look at patent collaboration we also see a similar pattern. 8 out of 13 regions are always a top-5 percentile region. 6 regions of Germany are always at the top-5 percentile list in patent collaboration. These regions are mostly located at the core of Europe, are populated (i.e., large human capital) and known to be investors in research and innovation early on which signifies the importance of starting levels. Starting level knowledge, infrastructure, human capital and geography may perhaps have given certain advantages to these regions so that they have a persistent network position. Most of the regions listed in the left panel are in 10-core countries that constituted the European Economic Community (EEC) and had a chance to participate in the first round of FP between 1984-1987.[5] It could be the case that such links in project collaborations are carried to collaboration in patents later on, which could also explain why there is persistence in top-5 percentile regions according to network degree.

Figure 4: Depicting network degree of FP project and patent collaborations in European NUTS2 regions

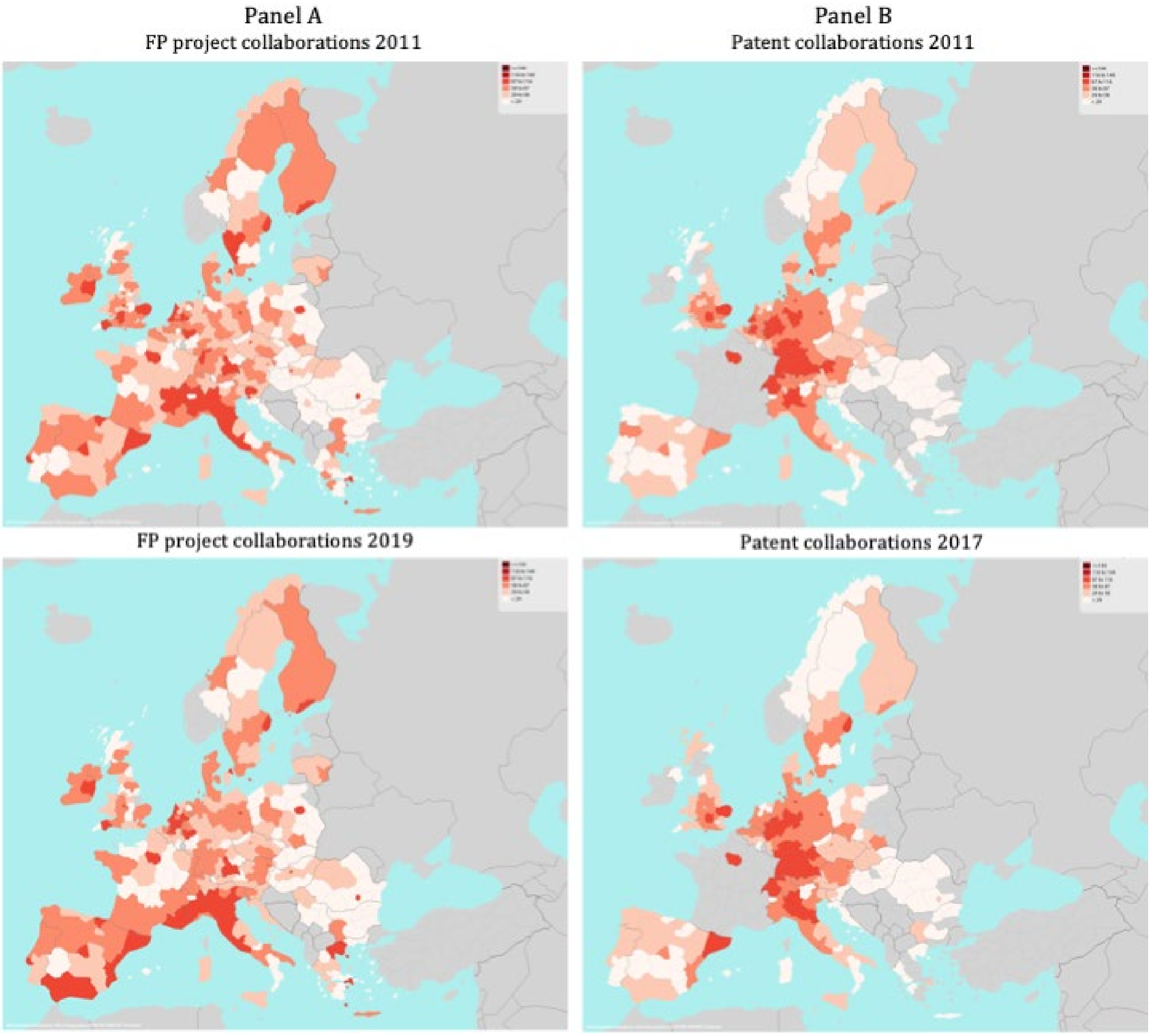


Figure 4 visualizes Table 7. Comparing panel A and panel B it is obvious that patent collaboration is a EU-core phenomenon. Darker red regions are the ones that have a high degree centrality where such regions are concentrated in south Germany, Switzerland and North of Italy. This pattern changes slightly over the years (2017 compared to 2011). In 2017 we also see regions from Spain, Finland and Sweden that are peripheral to Europe. When we look at Panel A, the FP project collaborations, the core and the

[5] EEC consisted of Belgium (BE), Denmark (DK), France (FR), Germany (DE), Greece (EL), Ireland (IE), Italy (IT), Luxembourg (LU), Netherlands (NL) and the United Kingdom (UK).

periphery are not as sharp as in Panel B. This could be because FPs already provide a platform for regions to collaborate and that one project collaboration may even connect 10 different regions. Tables 5 and 6 also show that degree centrality is much higher in FP project collaborations compared to patent collaborations. Comparing 2011 to 2019 in panel A of Figure 4 one can see peripheral regions that become better connected over the years. However, looking from the "number of connections" angle, degree centrality analysis tells us a story of persistence of top players.

Table 8 presents top-5 percentile regions according to betweenness centrality. This measure is affected by the network structure and the position of the regions that a particular region is connected to. Thus it shows the importance of a region in knowledge exchange within a network. High scores of betweenness centrality means that a region through its connections may act as a hub that connects unconnected regions. This position may help such regions to accumulate and utilize knowledge better than the others. Table 8 shows that the top 5 percentile region list according to betweenness centrality is much more heterogeneous compared to degree centrality. Regions listed in Table 7 are mostly present in Table 8 strengthening the discussion on persistence but we also see peripheral regions in the top-5 list such as Budapest (HU11), Warszawski stołeczny (PL91), Attica (EL30), Área Metropolitana de Lisboa (PT17) and Basque Community (ES51). Thus we can say that regarding the importance of positions of regions in a knowledge exchange network, persistence is a pattern but less pronounced and there are signs of emerging knowledge hubs at the periphery which may be taken as a sign of knowledge convergence.

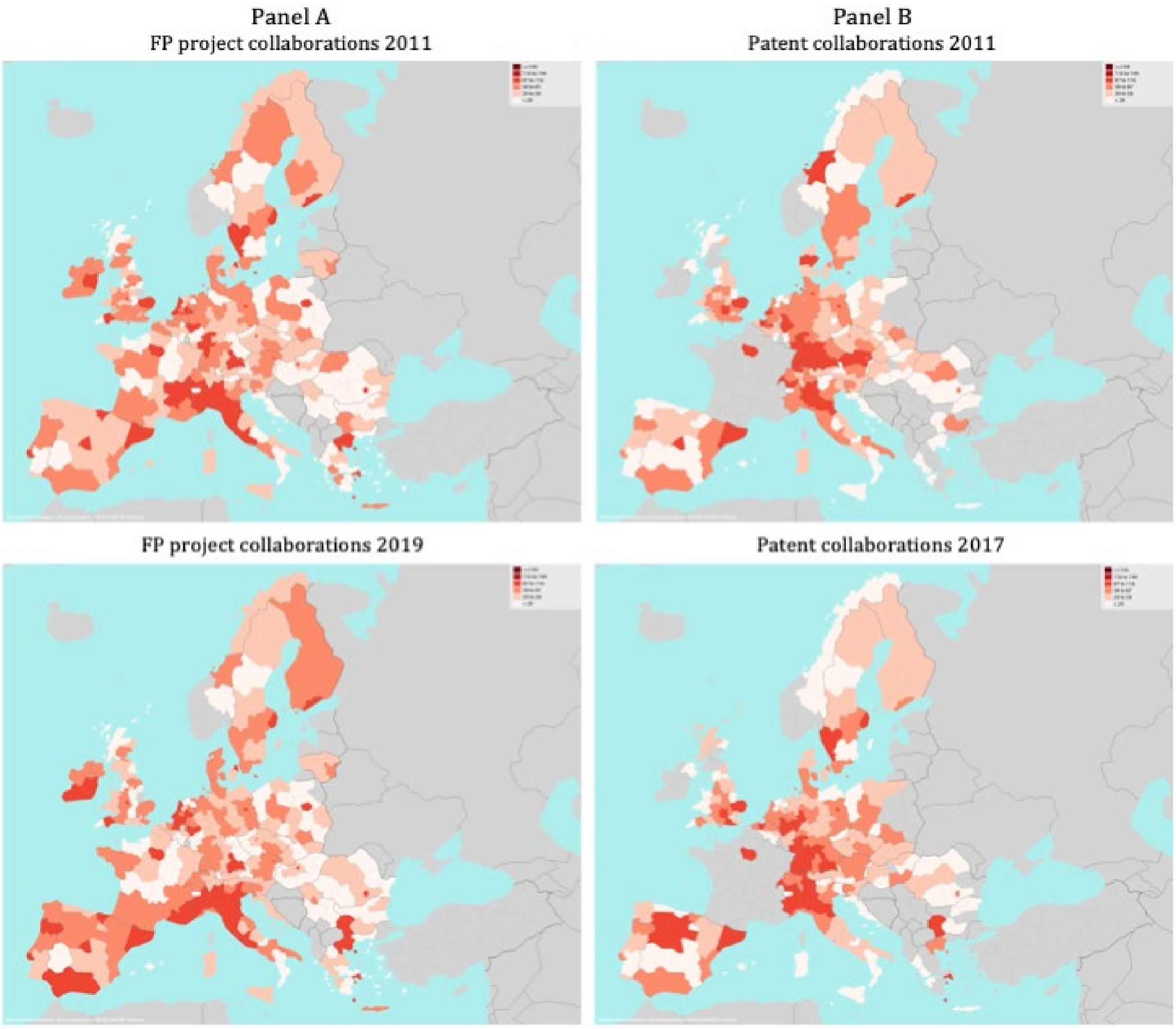
Table 8: Top 5 percentile NUTS 2 regions (according to network betweenness) - FP project collaborations vs. patent collaborations

FP project collaboration					Patent collaboration				
2011	2013	2015	2017	2019	2011	2013	2015	2017	2019
ITI4	ES30	ES30	ES30	FR10	DE21	DE21	DE21	DE21	DE11
EL30	BE10	FR10	BE10	ES30	DE71	FR10	FR10	DE11	DE21
FR10	FR10	ITI4	DEA2	NL33	ITC4	DEA2	DEA1	DE71	DEA1
BE10	ITI4	BE10	FR10	BE10	FR10	DE11	CH04	FR10	FR10
ES30	DE21	ES51	NL33	ITC4	DE11	ES30	DE71	DEA2	DE12
DE21	AT13	DE21	ITI4	UKI3	DE60	DE12	ITC4	DE12	ES51
UKI3	ES51	UKI3	DE21	ITI4	CH01	ITC4	UKH1	DEA1	DE71
ITC3	UKI3	ITC4	UKI3	FI1B	DE12	DE71	DE11	CH04	ITC4
ITC4	EL30	NL33	AT13	EL30	DEA2	DEB3	DEA2	ITC4	ITH5
ITI1	FI1B	DEA2	ITC4	ITH5	ITH5	DEA1	ES30	DE25	CH03
ES51	NL33	EL30	ES51	ES51	DEB3	CH01	DE12	ES51	NL41
NL33	ITH5	AT13	PT17	PT17	ES30	DEA5	DEB3	UKJ1	CH04
ITC1	ITC4	DE30	NL22	AT13	BE24	DE25	ES51	UKH1	
NL32	DEA2	NL22	EL30	DE21					
FI1B	ITC1	HU11	PL91	ES21					
DK01	HU11	DK01	FI1B	DEA2					

Comparing Figure 5 Panel B to Panel B of Figure 4 we see that in patent collaboration a core still exists but there are also regions from Spain, UK, Finland, Sweden and Greece that are similarly positioned in the patent collaboration network. Comparing 2011 to 2017 in Figure 5 Panel B one can observe that over the years some peripheral regions have increased their importance in the network. The existence of the core reflects persistence and the increased heterogeneity and emergent knowledge hubs in the periphery may reflect knowledge convergence. Figure 5 Panel A supports such findings. FP collaboration data is richer (includes more NUTS 2 regions) compared to patent collaborations and the emergence of peripheral knowledge hubs in Poland, Spain, Portugal, Ireland and Greece can better be observed.

This set of descriptive analysis shows that knowledge in Europe is produced in the core and there is persistence over the years. It may be difficult to tap into such a network but the findings also show that some peripheral regions may be considered as emerging knowledge hubs which may be taken as a sign of convergence.

Figure 5: Depicting network betweenness of FP project and patent collaborations in European NUTS2 regions



The above descriptive analysis can be extended in two ways. First, it could be the case that core or big cities of peripheral countries are better positioned in the knowledge exchange network but there is no overall pattern of knowledge convergence. To investigate this, simple scatter plots of difference of percentile ranks between 2019 and 2011 (for patent collaboration 2017 and 2011) and initial level of percentile rank can be used. Second, we can benefit from empirical economic growth literature where the variations in the change of Gross Domestic Product (GDP) are explained with a set of control variables and the initial level of GDP. A negative coefficient of initial level of GDP is taken as a sign of economic convergence. This idea can be utilized to assess whether there is knowledge convergence.

Figure 6: Knowledge convergence, FP project collaborations 2011-2019 vs. patent collaborations 2011-2017

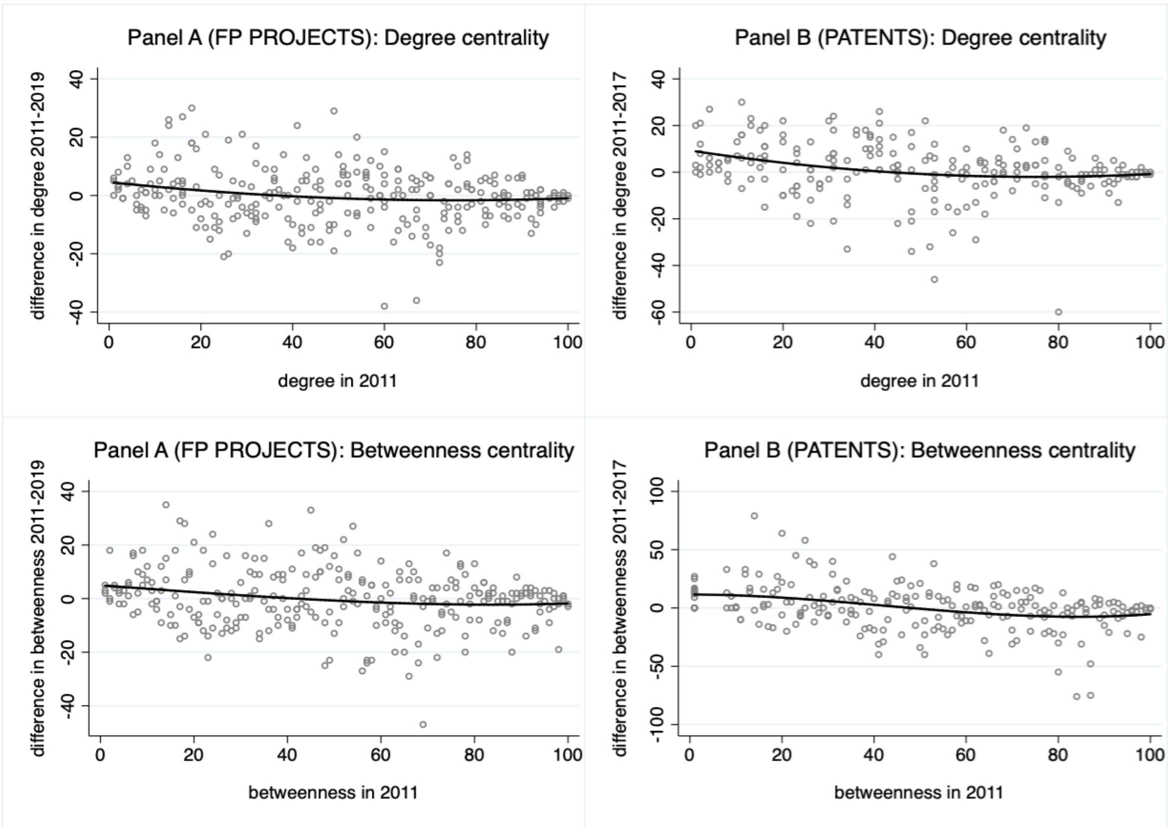


Figure 6 depicts simple correlations of initial levels and changes over for 2011-2019 FP project collaborations and for 2011-2017 patent collaborations. Panel A shows the scatter plot for degree centrality and betweenness centrality using FP project collaborations and Panel B shows the scatter plots for patent collaborations. To produce these figures, first percentile ranks of centrality values are computed. Higher percentiles represent more linkages (degree) and better position (betweenness) in the network. The difference of percentile ranks over a period (e.g., degree percentile rank 2019 - degree percentile rank 2011) may give clues about how regional position changes over the years. For instance, if the degree percentile rank is 50 in 2011 and 80 in 2019 it means that this particular region is better positioned in the network over time. The correlation of this change indicator and its initial level for different centrality measures ranges from -0.18 to -0.25 all of which are significant at the 1% level of significance. Thus the negative correlations depicted in Figure 6 are all statistically significant. Figure 6 also highlights the persistence story as can be understood from little variation after the 90th percentile (x axis) or similarly falling variations as one moves to higher percentiles. Both patent and FP project collaborations lead to knowledge accumulation. When regions enhance collaborations they not only occupy a better position in the knowledge network but also accumulate knowledge. Figure 6 shows that, on average, regions that are relatively poor in knowledge are catching-up regions that are rich in knowledge (indicating convergence). The level of knowledge in a region is approximated by its position in the knowledge network that is either based on FP project collaborations or patent collaborations.

Figure 7: Added variable plots of degree and betweenness in 2011. FP project collaborations 2011-2019 vs. patent collaborations 2011-2017

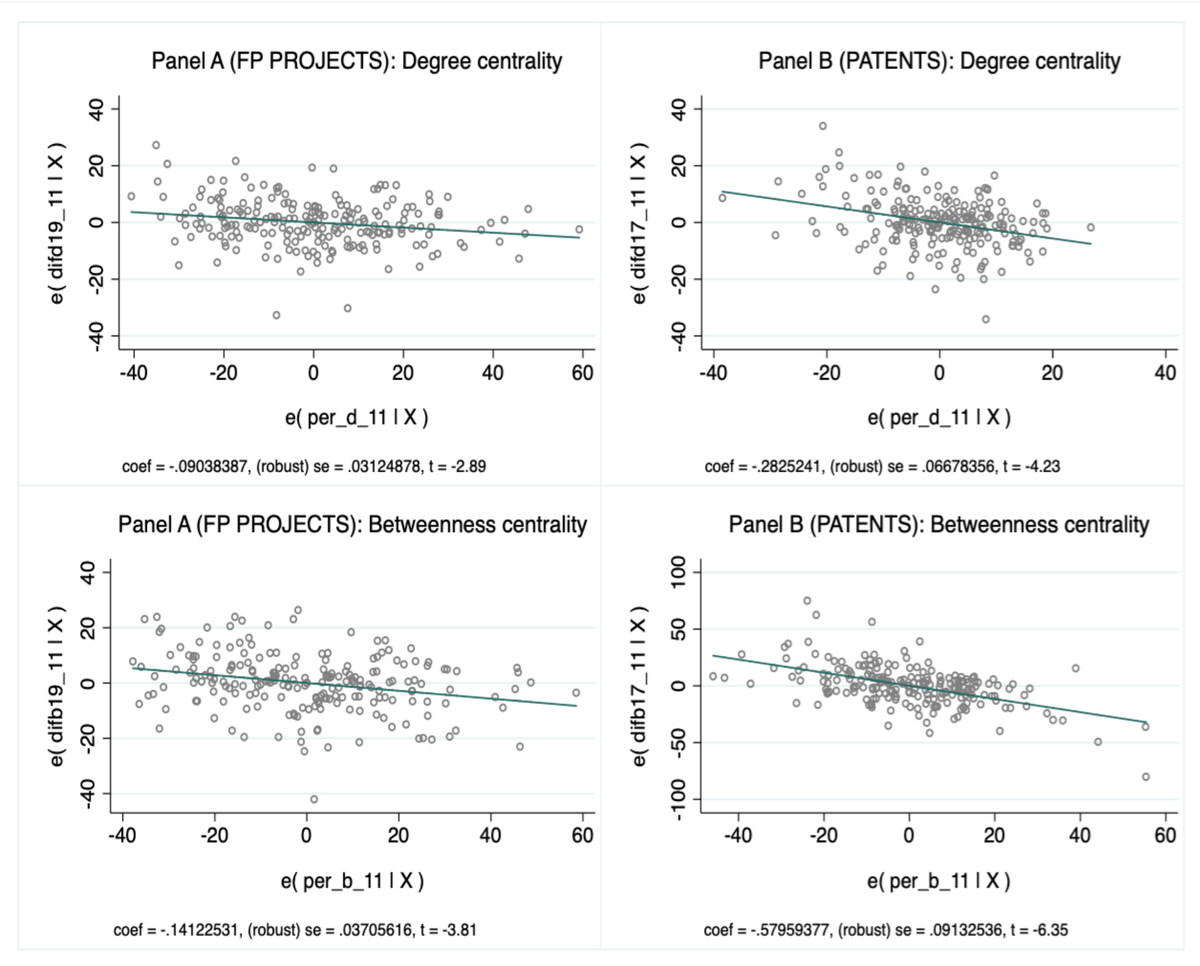


Figure 6 could be taken a step further with an OLS estimation where the variations in the difference in centrality rank over a period is explained by initial centrality rank, initial log population to control size effect, initial patent applications to control for initial knowledge stock and country fixed effects. The added variable plots of initial centrality rank using FP collaboration and patent collaboration data are presented in Figure 7. The negative coefficient of the difference of centrality rank 2019-2011 can be interpreted as evidence for knowledge convergence (similar to OLS regression of GDP growth on initial GDP and a set of controls). Panel A of Figure 7 shows the added variable plots of degree centrality and betweenness centrality that come from the FP project collaboration network. As can be seen, both coefficients are negative indicating convergence. In a similar manner, Panel B shows the added variable plots of degree centrality and betweenness centrality that come from patent collaborations. Compared to the FP project network, convergence has a stronger pattern in the patent network. In summary, knowledge exchanges within both the FP project and the patent network display a convergence pattern where regions that are less endowed with knowledge tend to catch up with knowledge-rich regions. While this finding is promising, it is still a correlation, it does not consider the impact of changes in the network structure and does not say much about cohesion.

4.3. SIENA results

SIENA results on FP project collaborations and patent collaborations are displayed in Table 9 and 10, respectively. Each table presents two panels where the top panel (network dynamics) shows the impact of regional characteristics (innovation and trust) network indicators (density, transitivity, # of links, strong tie) and control variables (distance to control for geography and employment in science and technology to control for size) on the network structure. The bottom panel shows the impact of network structure on regional characteristics. We expect network structure to affect regional characteristics to talk about cohesion. The analysis includes a total of 281 NUTS 2 regions. Jaccard coefficient of the simulation results in Table 9 shows the similarity between waves, and a value of 0.80 is sufficient, which is higher than the required level of 0.3.

Considering FP projects (Table 9), there are costs and benefits of establishing links to other regions. For instance, while a project partnership can bring new information and new network opportunities to the region; the burden of costs required to realize the project may be higher than the benefits. Density (outdegree) value is found to be negative and significant, meaning that nodes do not link randomly but are selective. This demonstrates that establishing and managing ties in FP projects are costly, as expected.

The results also indicate a significant coefficient for transitivity in networks. In other words, the probability of two regions forming a collaboration with each other depends on commonly known third parties, which has been shown to be the case in many knowledge networks in the literature (Gulati, 1999). Transitivity value is found to be positive and significant which shows that the nodes have a tendency towards clustering.

Table 9: Siena results of FP project collaborations 2011-2019

Network dynamics: the impact of regional characteristics on network structure			
	coeff.	s.e.	interpretation
network (period 1)	4.905	0.388	On average each region is selected as a partner 4.9 times between p1 and p2.
network (period 2)	6.291	0.646	See the above interpretation.
network (period 3)	6.245	0.509	See the above interpretation.
network (period 4)	5.067	0.353	See the above interpretation.
Density (Outdegree)	-7.609***	0.202	Nodes do not link randomly, they are selective. Establishing and managing ties are costly.
Transitivity	2.420***	0.098	Tendency towards transitivity. Evidence of clustering.
# of links (node)	0.033***	0.001	The more links a node has the more it sets up.
Strong tie	2.383***	0.078	Previous collaboration drives further collaboration.
Distance	-0.096**	0.039	Distance between regions has a negative effect on collaboration.
Innovation ego	0.0006	0.052	Being highly innovative has no significant effect on collaboration.
Innovation similarity	0.313***	0.105	Collaboration tends to be high among regions with similar innovation levels.
Trust ego	0.187**	0.075	Being a high trust region significantly affects collaboration.
Trust similarity	-0.342	0.259	No significant effect of collaboration among regions with similar levels of general trust.
Employment ego	0.113**	0.050	Being rich in terms of human resources in S&T significantly affects collaboration.
Employment similarity	0.024	0.113	No significant effect of collaboration among regions with similar levels of Human Resources in S&T
Behaviour dynamics: the impact of network structure on regional characteristics			
rate (period 1, innov.)	0.013	0.009	Indicates the average number of change opportunities regarding increasing innovation level.
rate (period 2, innov.)	0.374	0.057	See the above interpretation.
rate (period 3, innov.)	0.696	0.081	See the above interpretation.
rate (period 4, innov.)	0.490	0.067	See the above interpretation.
Innov. lineary shape	-0.389**	0.198	There is steady decrease in innovation level over time
Innov. quadratic shape	0.050	0.050	No statistically significant effect.
Innov. degree	0.009***	0.004	The more active the region is (i.e., collaboration, links) The higher its innovation level.
Innov. average alter	0.415	0.302	No effect of neighbours on the region.
rate (period 1, trust)	0.709	0.113	Indicates the average number of change opportunities regarding increasing general trust level.
rate (period 2, trust)	0.999	0.158	See the above interpretation.
rate (period 3, trust)	1.397	0.206	See the above interpretation.
rate (period 4, trust)	1.209	0.163	See the above interpretation.
trust linear shape	0.040	0.075	No statistically significant effect.
trust quadratic shape	-0.378***	0.065	Negative feedback indicates a self-correcting mechanism. The push for increasing trust for regions that already have high trust becomes smaller.
trust degree	0.002	0.003	Region's activeness in collaboration has no impact on the region.
trust average alter	0.675	0.524	No effect of neighbours on the region.

Note: The network and rate variables indicate expected frequencies, between successive periods, with which nodes (regions) get the opportunity to change a network tie. Though we report the standard errors of these variables, testing that they are zero is meaningless because if the coefficients are zero there is no change in the network at all (Snijders et al., 2010).

One of the activities conducted in all FP projects is related to the visibility of regions. Visibility activities are carried out to promote European Commission's support and to increase the visibility of the activities performed and results obtained throughout a project. Such promotions render project partners ("project partner" is used as a generic term without discriminating between coordinator and partner) visible to the institutions/agencies outside the project. As a result, partners that are involved in many projects can be more visible, and attract more partners in the future. In the literature, such a "rich get richer" effect has been studied under the concept of preferential attachment (Barabasi and Albert, 1999). The number of prior links that a region has is a factor facilitating its establishing links later on. Number of links (node) value is found to be positive and significant; in other words, the number of links a node has makes a positive impact on establishing new links.

A collaboration between two regions in a project creates a potential to establish collaborations in future projects. Through collaboration, parties get to know each other, develop a common language, which facilitates their working together in the subsequent projects. The strength of ties not only supports the quality of information to be transferred between nodes, but also affects the status of the relationship between the parties. Repeated interactions lead to the strengthening of ties in between. This repetition will also act as a factor that influences parties to naturally prefer each other in other projects. Strong tie value is found to be positive and significant, meaning that trust is a factor that has a positive impact on establishing links.

Geographical distance between two regions is another factor that may influence networks. Project collaborations between regions that are distant may reflect differences in cultural attributes, priorities and practices of business conduct. For example, a Mediterranean and a Baltic coastal region may drastically differ from each other in terms of project related practices and institutions. Distant geographies may also have different languages which may act as a barrier. In addition, distance may make it more difficult to collaborate on a more practical level. Such differences led by geographic distance are expected to have a negative impact on collaboration. Indeed, the results reveal that the coefficient of geographical distance value is negative and statistically significant which indicates that geographic distance has a negative impact on collaboration.

A region with a high innovation level demonstrates that it is successful in producing new knowledge, as well as absorbing incoming knowledge. It is expected that a high innovation level would be an attractive factor for establishing links; however, this hypothesis is not verified. While it could probably be attractive for a low innovation region to link with a high innovation region, the opposite may not hold (*i.e.* innovation ego value is positive but insignificant). On the other hand, the results reveal the significant positive effect of innovation similarity on network structure. Stated differently, regions with similar innovation levels are more likely to establish links with each other.

In terms of the trust variable, results indicate a positive and significant coefficient. In other words, the regions with high degrees of trust are engaged in more collaborations. The measure of trust here can be taken as an indicator of openness of a region to collaborations. The similarity in trust values between two regions has no significant impact on networks, according to the results.

Another variable that is included is the human capital of regions. Regions' having well-trained human resources in Science and Technology (S&T) is critical both in their production of new knowledge and learning capabilities. As the quality and number of human resources of the regions increases, they would naturally become more attractive for other regions. Indeed employment ego value was found to be positive and significant. On the other hand, although a positive tendency is observed towards establishing links with those regions with a similar level of human resources in S&T the results are insignificant.

The impacts of innovation and trust, which were defined as behaviours, as well as node characteristics on the shaping of the FP project network are also examined. First, the linear and quadratic shape of the objective function is checked. Innovation linear shape is found to be negative and significant, which points at the gradual decrease in the innovation value from the first wave to the last wave. Quadratic shape, on the other hand, show that the node with a high innovation value in period 1 has a higher innovation value in period 2; however, the results are insignificant. Two indicators related to innovation as behaviour are innovation degree and innovation average alter. Innovation degree is positive and significant, which means that NUTS regions with a higher degree (more "active" actors) have a stronger tendency toward high innovation, thus being in the network and collaborating tend to increase innovation level. On the other hand, despite a positive innovation average alter coefficient, which shows the tendency of nodes to reach to an innovation level similar to the nodes they established links with (influence) the results are insignificant. The story is a bit different for the second behaviour variable, generalized trust. Trust linear shape is statistically insignificant. Quadratic shape, on the other hand, is significant and shows that those with a lower trust value in 2011 have a higher trust value in 2019.

Table 10: Siena results of Patent collaborations (inventor network) 2011-2017

Network dynamics: the impact of regional characteristics on network structure			
	coeff.	s.e.	interpretation
network (period 1)	7.918***	0.787	On average each region is selected as a partner 7.9 times between p1 and p2.
network (period 2)	7.018***	0.698	See the above interpretation.
network (period 3)	7.321***	0.767	See the above interpretation.
Density	-4.434***	0.197	Nodes do not link randomly, they are selective. Establishing and managing ties are costly.
Transitivity	1.387***	0.082	Tendency towards transitivity. Evidence of clustering.
# of links (node)	-0.031***	0.009	The more links a node has the less it sets up.
Strong tie	1.174***	0.073	Previous collaboration drives further collaboration.
Distance	-0.862***	0.135	Distance between regions has a negative effect on collaboration.
Innovation ego	-0.001	0.063	Being highly innovative has no significant effect on collaboration.
Innovation similarity	0.627**	0.294	Collaboration in patents tend to be high among regions with similar innovation levels.
Trust ego	-0.051	0.123	Being a high trust region has no significant effect on collaboration.
Trust similarity	2.951***	1.005	Collaboration tends to be high among regions with similar trust levels.
Employment ego	0.257***	0.093	Being rich in terms of human resources in S&T significantly affects collaboration.
Employment similarity	0.057	0.166	No significant effect of collaboration among regions with similar levels of human resources in S&T
Behaviour dynamics: the impact of network structure on regional characteristics			
rate (period 1, innov.)	0.012	0.009	Indicates the average number of change opportunities regarding increasing innovation level.
rate (period 2, innov.)	0.363	0.074	See the above interpretation.
rate (period 3, innov.)	0.673	0.099	See the above interpretation.
Innov. linear shape	-0.414	0.387	No statistically significant effect.
Innov. quadratic shape	-0.243	0.255	No statistically significant effect.
Innov. degree	-0.053	0.060	Own activity of the region is (i.e., collaboration, links) has no effect on innovation level.
Innov. average alter	1.545	1.153	No effect of neighbours on the region.
rate (period 1, trust)	0.963	0.214	Indicates the average number of change opportunities regarding increasing innovation level.
rate (period 2, trust)	1.310	0.256	See the above interpretation.
rate (period 3, trust)	1.447	0.284	See the above interpretation.
trust linear shape	0.321	0.302	No statistically significant effect.
trust quadratic shape	-1.701*	0.946	Negative feedback indicates a self-correcting mechanism. The push for increasing trust for regions that already have high trust becomes smaller.
trust degree	0.033	0.029	Own activity of the region is (i.e., collaboration, links) has no effect on trust.
trust average alter	5.196	1.500	No effect of neighbours on the region.

Note: The network and rate variables indicate expected frequencies, between successive periods, with which nodes (regions) get the opportunity to change a network tie. Though we report the standard errors of these variables, testing that they are zero is meaningless because if the coefficients are zero there is no change in the network at all (Snijders et al., 2010).

Patent network (Table 10) displays similarities to FP network (Table 9) in certain aspects. Even though significance levels differ, density, transitivity, tie strength, distance, innovation ego and innovation similarity, employment ego, and employment similarity variables have similar effects in both networks. As in the FP networks, nodes behave selectively rather than randomly when establishing links (negative and significant coefficient of density). Transitivity coefficient is also positive and significant for patent network, which indicates that the likelihood of collaboration between two regions increases with common partners. However, as different from FP networks, we do not detect a preferential attachment mechanism; on the contrary the negative coefficient of number of links indicates that the more links a region has, the less links it attracts. This might be because of a certain saturation level in the total number of patent collaborations in a region. Likelihood of collaboration increases with repeated ties between two regions, as revealed by the positive and significant coefficient of strong tie. On the other hand, the tendency to cooperate decreases with geographical distance between two regions as indicated by the negative and significant coefficient of distance. Rather than the innovation level of the node, the similarity of the innovation levels between two regions (positive and significant coefficient of innovation similarity) impacts collaboration. High levels of employment in S&T shape the network (*i.e.*, establishing or terminating links). Interestingly, the effect of the trust variable is different from the FP network analysis. The results indicate a significant positive effect on networks of similarity in trust levels between two regions. In other words, regions with high trust values tend to collaborate more with high trust regions.

When behaviour variables are analyzed, innovation value decreases over time from period 1 to period 3. However results are insignificant, in other words innovation is linear-shaped. Although the value of the quadratic shape is found to be insignificant; over time the value of the nodes with high innovation value decreases, while the value of those with low innovation values increases. In other words, the innovation value does not reinforce itself. In the patent network, neither innovation degree nor innovation average alter values are found to be significant. That is, NUTS with a higher degree (more 'active' actors) do not have the tendency to reach a higher innovation level; nodes, also, do not tend to reach innovation levels similar to the nodes they established links with (thus no influence effect as well). Regarding trust, while linear shape is insignificant, quadratic shape coefficient is significant; albeit with a relatively low significance level. Over time, the value of those with high trust values decreases while that of nodes with lower values increase. Thus there is a self-correcting mechanism. Trust degree and trust average alter are insignificant in the patent network just as in the FP project network.

Comparing Table 9 and 10 we can reach two conclusions regarding knowledge cohesion: (i) it seems that both in the patent and FP project network, regions tend to collaborate with regions that are similar to themselves. In terms of network structure research and invention networks resemble and there is more evidence toward convergence (ii) the impact of network structure on regional characteristics is mostly statistically insignificant especially in the patent network. We take these as evidence against knowledge cohesion.

These results reveal some differences between FP project and patent networks in terms of the mechanisms that drive their evolution. Above (section 4.1 and 4.2) it is explained that we observe some convergence between regions in FP project network. These results are better explained when we analyze SIENA output. It seems that there are self reinforcing mechanisms at work in FP project network; strong network transitivity and tendency towards closure, strong influence of strong ties between regions on network evolution are some of the indicators of this mechanism. On the other hand, comparison between the two networks reveals that network evolution in patent network is more sensitive to some of the variables, compared to FP project network. An in depth exploration of these variables will enable gaining some insights.

Firstly, transitivity mechanism in networks is nearly two times higher in FP project networks than in patent network.[6] Secondly, patent networks are more sensitive to distance between regions. Geographically proximate regions are significantly more likely to collaborate as compared to FP project networks. Third, the effect of strong ties on network evolution is stronger in FP project networks, compared to patent networks. This means that repeated collaborations between regions have a higher role in driving further collaboration in FP project networks. Fourth, the innovation similarity between two regions plays a more important role in patent networks, compared to FP project networks, as revealed by the value of the coefficients of innovation similarity variable. Finally, regional human capital plays a more important role in patent networks than FP project networks.

These factors imply that patent network evolution is *more* sensitive to regional attributes than FP project network: distance between regions, employment levels, and innovation similarity between regions. At the same time, patent networks are *less* sensitive to variables related with the past networks than FP project networks: common collaboration partners (transitivity), effect of past collaborations and density have comparably less role in network evolution in patent networks. These results may indicate that patent collaborations are more sensitive to market-based changes in the environment. In other words, for patent networks, environmental or regional characteristics are more important than collaborations in the past. On the contrary for FP project network, past networks impose a stronger self reinforcing mechanism, as revealed by the strong role of past networks on the evolution of FP networks.

In terms of knowledge convergence and cohesion, the results indicate that in both cases there are signs of convergence. But for patent networks, there seems to be stronger convergence between similar regions which are closer to each other physically, which can

[6] In the RSIENA Manual, Ripley et al (2020), section 8.5 (p. 100) discusses comparison and testing between two different networks and shows that if the network model and the observations are the same coefficients of two different networks can be compared.

increase the discrepancies between all the regions in the long run. This possible fragmentation is likely to inhibit knowledge cohesion in the long run. For the FP project network, on the other hand, the strong role of past networks on network evolution prepares a more suitable ground for knowledge cohesion between regions. However in the period we analyse we do not find strong evidence for cohesion.

Table 11: Comparison of two networks' evolution mechanisms

What drives network evolution?	***High effect***	***Low effect***
<u>Regional attributes</u> ● Distance between regions, ● Human capital ● Innovation level similarity between regions	Patent networks	FP networks
<u>Past networks</u> ● Strength of ties ● Density ● Transitivity	FP networks	Patent networks

5. Conclusion

In this research we investigate knowledge convergence and knowledge cohesion in the EU by analysing FP project collaborations and patent inventor collaborations over the 2011-2019 period using network analysis, but especially SIENA that allows to analyse the impact of network structure on regional characteristics. Our results can be summarized in three steps.

First, looking at the SNA results there are similarities in the evolution of networks between the science and invention networks as well as divergences. In the case of science networks (FP project) regions tend to converge to each other, as the network statistics reveal. Reduced geodesic distances, higher network density and average degrees of nodes, and higher clustering coefficients can be interpreted as signs of knowledge convergence. These findings reveal that peripheral regions are increasingly connecting to the core through project collaborations which we assume to result in knowledge accumulation in the periphery to a certain extent. While we observe similar patterns in the patent invention networks initially, there are signs of loosening of networks after 2015. This finding may mean that the two networks are possibly driven by different mechanisms of evolution in time. In summary, our results contribute to the collaboration-induced knowledge diffusion literature, specifically Balland, Boschma and Ravet (2019). We extend their findings at the country level to the NUTS2 regional level. We also provide a first analysis of knowledge convergence as conceptualized in section 2.

Second, looking at the descriptive analysis using degree centrality and betweenness centrality we find that although there is strong persistence among the knowledge hubs in the core (i.e., top knowledge-rich regions in 2011 are still knowledge-rich in 2019) there is a certain degree of catching-up. A simple OLS estimation shows that even after controlling for size, initial knowledge stock and country fixed effects, there is a negative correlation between initial network statistics (the level) and the difference over the 2011-2019 period indicating catch-up. Especially in terms of the importance of the position of regions in the information exchange (i.e., betweenness centrality) we observe that over the years, some peripheral regions of core countries such as Germany and capital cities and/or large cities of peripheral countries tend to obtain a better position in the FP project network. For instance, the diversity in top 5 percentile regions according to network degree and betweenness increases in research networks but is fairly stable in patent networks. These findings suggest that there may be a certain degree of knowledge convergence among the NUTS2 regions of the EU over the 2011-2019 period.

Finally, to analyse the existence of knowledge cohesion we utilized SIENA to see what factors affect the evolution of science and invention networks and whether structural changes in the network affect regional characteristics. Science network and invention network display similarities in many aspects. Regions behave selectively when establishing links, the likelihood of collaboration increases with common partners and with repeated ties between two regions. We also find that the tendency to cooperate decreases with geographical distance. Such findings signal convergence. We further find that in both the science and invention networks regions tend to collaborate with regions that are similar to themselves in terms of innovation level and the network structure do not have statistically significant effects on regional characteristics. These findings led us to argue that collaborations in science and invention networks tend to form knowledge convergence but not knowledge cohesion.

We contribute to the literature by first suggesting a rather novel way to investigate (or measure) cohesion. To our knowledge this research is one of the first in economics and management studies that utilized SIENA. Second, we strictly differentiate between knowledge convergence and knowledge cohesion by conceptualizing knowledge cohesion. For instance, to investigate economic cohesion researchers mostly use econometrics and provide analysis of changes in economic outcome indicators such as GDP (e.g., Sala-i-Martin, 1996; Ederveen, de Groot and Nahuis, 2006; Becker, Egger and Ehrlich, 2010; Pellegrini et al., 2013; Fiaschi, Lavezzi and Parenti, 2018). In essence, such methods and measurement provide evidence for convergence but not for cohesion. Third, we used two different data sources and analysed science and invention networks separately. Reaching similar results in two different networks increase the robustness of our findings.

This research has several limitations that could be improved in future research. One may argue that a 10-year period is too short to analyse knowledge cohesion, thus our main finding that there is knowledge convergence but not cohesion among regions may be expected. While we sympathise with this idea, data availability at the regional level (innovation, trust, employment) was the sole reason why we base the analysis on the 2011-2019 period. Unfortunately, the regional innovation scoreboard data is only available from 2011 onwards. Data availability is a problem especially when the unit of analysis is NUTS2 regions. The analysis period could be extended by selecting a different set of regional indicators, for instance using GDP instead of innovation. Another problem in our analysis is we cannot form a directed network. Due to the nature of data it is not possible to know the direction of knowledge transfer but apart from that in section 2 we argue that even though there are knowledge stock differences between regions, both regions can learn and obtain new knowledge from a collaboration. Finally, in PATSTAT data, matching address information to NUTS2 codes was a labour intensive cumbersome task. The better this is done the more data will be available.

5.1. Policy implications

In this sub-section, we will discuss the possible policy implications of the results obtained in the project. However, in order to discuss them, we need a policy framework in the context of EU policies with regard to the scope of the data used. This framework is provided by the evolution of EU cohesion policy and a relatively younger policy tool of smart specialization. Thus, we first provide a snapshot on these issues and link this discussion with the findings of our research. This study is an attempt to bridge the repercussions of EU cohesion policy with the observed data and some stylized facts. Although our results find evidence in favor of convergence, regions have an inclination to collaborate with regions that are similar to themselves in terms of innovation level and general trust. This finding especially questions the success of longstanding EU cohesion policy.

Policy aims and tools to reach the targets of any policy prescription can be observed at various levels of abstraction and implementation. Nowadays, a multi-dimensional approach in designing public policies especially in the era of rising complexity in inter- and intra-organizational relations seems to be necessary at regional, national and supranational levels. A failure to integrate this multi-dimensional insight in policy prescriptions would not only fail to solve the existing problems but also deepens them. A multi-dimensional approach helps decision-makers in explaining the embedded complexity and overcoming barriers in designing successful public policies. The nested relations at each dimension should be examined and the feedback loops are carefully defined for the evolution of policies and co-evolution of the policy environment.

European Union (EU) cohesion policy is a cross-cutting arena including a complex web of relations. It is one of the most important policies of the EU and historically one of the most financially significant with its considerable share of the EU budget. In 2013, we observed a milestone with a set of significant changes by building a new policy direction and in conformity with the motto of the Europe 2020 strategy for smart, sustainable and inclusive growth. This date almost perfectly coincides with our sample of observations in the analysis of this paper.

In the context of EU policy making, EU's growth strategy, called Europe 2020, has five assertive objects on employment, innovation, education, social inclusion and climate/energy. ERDF for the 2014-2020 has introduced an ex-ante condition which requires all EU member states on national and regional levels to have a Research and Innovation Strategy for Smart Specialisation (RIS3) for the process of entrepreneurial discovery before their operational programmes are approved (EC, 2014). RIS3 strategy is the most important target of the EU Cohesion policy. Among others, one of the most important concepts in this period is smart specialization. The international debate on RIS3 started in 2009 with the Knowledge Economist Policy Brief on "Smart Specialisation –The Concept" (EC, 2009) which was authored by three pioneering economists, Dominique Foray, Paul David and Bronwyn Hall (Capello and Lenzi, 2016). The report emphasizes the importance of specialisation on R&D and innovation at the regional level (Erdil and Çetin, 2019:215). Their plan on RIS3 stands on one very simple idea, "entrepreneurial process of discovery" which entrepreneurs act as leaders for future specialisation of regions. The logic of RIS3 could be simplified with the framework of "General Purpose Technologies (GPTs)". GPTs such as steam engine, electricity and computers have the capacity to change economies by affecting multi-sectors (Foray et al., 2009). Foray et al. (2011) appreciate the success of the concept of smart specialisation but warn the policy makers and academics about the fact that there is no sound base of empirical work in favour of the concept and underlines the fact that there is a growing gap between policy making and the theory. Smart specialisation involves discovering the unique and original components of the regional knowledge base and a policy-making process in conformity with these characteristics. Thus, there should be greater empirical work on these issues as well as the impact of the policies. The concept is a very young one and still needs empirical verification (Erdil and Çetin, 2019:215).

The specialisation of one sector in a region that is economically valuable and important is not certainly a smart specialisation. The smart specialization strategy needs a policy design that focuses on the encouragement of entrepreneurs (Erdil and Çetin, 2019:215). Foray (2015) especially warns that the concept of specialisation is different than the usage in localization (agglomeration) economies. Regional concentration of knowledge and competences form the "specialisation" notion in RIS3 instead of the relative concentration of one industry in a country. "Smart" notion is smart because RIS3 connects all the actors (such as entrepreneurs, government and local actors) and encourages the regions to be ambitious but realistic (Foray et al., 2012).

Europe's RIS3 strategy is inspired from four leading elements which depends on the past experiences and it needs general transformation of these principles which are summarized as "four Cs": Choices and Critical Mass, Competitive Advantage, Connectivity and Clusters, and Collaborative Leadership (Erdil and Çetin, 2019:216). RIS3 guide

(Foray et al., 2012) clearly defines the step-by-step process of designing a smart specialization strategy in six steps:

- Analysis of the Regional Context and Potential for Innovation: The analysis contains the analysis of the regional assets, linkages to the rest of the world and dynamics of the entrepreneur environment;
- Governance. Ensuring participation and ownership: Due to the collaboration of different stakeholders, it is necessary to have "boundary spanners" who should harmonize and moderate the RIS3 process;
- Elaboration of an Overall Vision for the Future of the Region: As the RIS3 is a long-term project, it is necessary to keep all the stakeholders in the process;
- Identification of Priorities: Potential areas of smart specialisation should be identified to trigger the regional potential;
- Definition of coherent policy mix, roadmaps and action plan;
- Integration of monitoring and evaluation mechanisms.

Capello and Kroll (2016) underline the fragilities of the RIS3 approach from theory to practice. They note that the new paradigm is a shift in the cohesion policy which promotes endogenous development, continuous innovation and a growth perspective (Capello and Kroll, 2016: 10). They advocate that the design of policies in the context of RIS3 should consider cohesion and competition goals simultaneously and take RIS3 as a good starting point for the development of cohesion policy. The success of RIS3 depends on its potential to transform knowledge and innovation into local development by using regionally untapped resources (Erdil and Çetin, 2019:216). In fact, Europe 2020 Strategy is the counter-move of the Commission to the failure of the Lisbon Strategy and its subsequent 2005 revision (Budd, 2013). The impacts of the global financial crisis of 2008 had still been felt during the implementation of this strategy. The conceptual framework used in our research suggests that excellence in technology-generating research is not automatically converted into commercial success. Deriving the economic impact from technology and innovation depends on dynamic interactive processes involving individuals, firms and institutions which absorb, apply and diffuse knowledge. Therefore, a broad set of framework conditions should exist for the optimal impact of innovation processes.

The most striking target of smart specialization policies is to perfectly construct complementary relations between excellence-based and place-based policies as a response to the linear research-based approach and corresponding European paradox. Although the findings of our analysis present partial evidence in favor of the success of smart specialization policies, the ultimate target seems not to be reached during the implementation. The synergy between excellence-based and place-based policies has seemed to be created at some extent in inter-regional level for collaboration towards knowledge convergence yet it seems to fail in the context of knowledge cohesion because of the problematic tension between excellence-based and place-based policies especially at the national level. While scientific excellence-based policy may not seem to be a high priority for every region, place-based specialization, innovation policy and knowledge cohesion are. This tension resembles itself as a barrier for a desired level of knowledge cohesion at the regional level.

The findings of selective nodes and transitivity as evidence of clustering drove knowledge convergence towards maturity. On the other hand, the co-evolution of both science and patent networks with a path-dependent history in which previous collaboration drives further collaboration may also be treated as an evidence for knowledge cohesion. Moreover, compared to the FP project network, convergence has a stronger pattern in the patent network in which collaboration in patents tend to be high among regions with similar innovation levels. This finding also underlines the above-mentioned tension for two different sets of policies. The actors seem to be more successful in the activities towards commercialization and mitigating the European paradox. However, the surprising finding of the insignificant effect of being highly innovative on collaboration verifies our presumption of the unsuccessful implementation for creating synergy between excellence-based and place-based policies towards knowledge cohesion. Another surprising finding is the irrelevance of being a high trust region in collaboration activities. However, collaboration tends to occur more among regions with similar trust levels. Although these findings seem to be surprising at the first instance, they constitute evidence in favor of knowledge convergence but not knowledge cohesion.

In conclusion, the empirical findings of the project are in conformity with the taxonomy presented in Figure 2. We end up with relatively few regions in “knowledge space II” as compared to “knowledge space III”. However, this does not mean a total failure of smart specialization policies. The implementation process of these policies suffers from the aforementioned problems and tensions, namely multi-dimensional coordination and agency problems. The next programming period is planned to fuel the innovation activities to mitigate the existing problems. For the next programming period of 2021-2027, the Commission proposed to modernize the Cohesion Policy in May 2018. The proposed Cohesion policy budget is approximately €331 billion for the 2021-2027 period as compared to €374 billion for the 2014-2020 period. The proposal can be evaluated as aspiring for the EU yet as pragmatic given the budgetary problems related to Brexit. It can be claimed that it is not myopic in the sense that its future orientation with the pressures of increase in innovativeness and underlining a strong commitment to solidarity. However, COVID-19 pandemic has created extra pressures for the implementation of the policy in the next programming period. The initial reaction of the Commission to the pandemic is a new instrument called the REACT-EU (Recovery Assistance for Cohesion and the Territories of Europe) package. REACT-EU will add fresh additional resources amounting to €47.5 billion to existing cohesion policy programmes. The real building block of the smart specialization policy was the entrepreneurial process of discovery. The success of the entrepreneurial process of discovery depends on involvement of stakeholders with past practices to realize overarching targets of knowledge cohesion. Nevertheless, this success is highly correlated with the stakeholder selection in a wider range of complementary assets. As a final word, the EU is in urgent need of a new generation regional innovation policy that does not ignore the co-evolution of the systems at regional, national and supranational levels with a multi-dimensional perspective towards knowledge cohesion that further ensures social and economic cohesion.

References

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development*, 59(1), 87–96.

Balland, P-A, Boschma, R. and Ravet, J. (2019). Network dynamics in collaborative research in the EU, 2003-2017. European Planning Studies, 27(9), 1811-1837.

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286, 509-512.

Becker, S.O., Egger, P.H., and von Ehrlich, M. (2010). Going NUTS: The effect of EU Structural Funds on regional performance, *Journal of Public Economics*, 94(9-10), 578-590.

Bergman, E.M. and Maier, G. (2009). Network central: regional positioning for innovative advantage. *Annals of Regional Science*, 43(3), 615-644.

Bouayad-Agha, S., Turpin, N. and Vedrine, L. (2013). Fostering the Development of European Regions: A Spatial Dynamic Panel Data Analysis of the Impact of Cohesion Policy, *Regional Studies*, 47(9), 1573-1593.

Boumans, D. and Ferry, M. (2019). Networks and efficient policy implementation: Insights from cohesion policy, *Evaluation*, 25(4), 411-429.

Brown, J.S. and Duguid, P. (1991). Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation. *Organization Science*, 2(1), 40–57.

Brown, A.L. and Campione, J.C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles, and systems. In R. Glaser (Ed.), *Innovations in learning: New environments for education* (pp. 289-325). Mahwah, NJ: Erlbaum.

Budd, L. (2013). Europe 2020: A Strategy in Search of a Regional Policy Rationale? *Policy Studies* 34(3), 274–290.

Capello, R. and Lenzi, C. (2016). Persistence in regional learning paradigms and trajectories: Consequences for innovation policy design. *European Planning Studies*, 24(9), 1587–1604.

Capello, R. and Kroll, H. (2016). From theory to practice in smart specialization strategy: Emerging limits and possible future trajectories. *European Planning Studies*, 24(8), 1393–1406.

Cheadle, J.E. and Goosby, B.J. (2012). The Small-school Friendship Dynamics of Adolescent Depressive Symptoms, *Society and Mental Health*, 2(2), 99-119.

Crescenzi, R. and Giua, M. (2016). The EU Cohesion Policy in context: Does a bottom-up approach work in all regions? *Environment and Planning A*, 48(11), 2340-2357

Dall'Erba, S. and Fang, F. (2017). Meta-analysis of the impact of European Union Structural Funds on regional growth. *Regional Studies*, 51(6), 822–832.

De Noni, I., Orsi, L. and Belussi, F. (2018). The role of collaborative networks in supporting the innovation performances of lagging-behind European regions. *Research Policy*, 47, 1-13

Di Caro, P. and Fratesi, U. (2021). One policy, different effects: Estimating the region-specific impacts of EU cohesion policy, *Journal of Regional Science*, forthcoming.

Di Cataldo, M., Monastiriotis, V. and Rodriguez-Pose, A. (2020). How 'Smart' Are Smart Specialization Strategies?, *JCMS-Journal of Common Market Studies*, forthcoming.

Dretske, F.I. (1981). *Knowledge and the Flow of Information*. Massachusetts: The MIT Press.

Ederveen, S., de Groot, H.L.F. and Nahuis, R. (2006). Fertile soil for structural funds? A panel data analysis of the conditional effectiveness of European cohesion policy, *Kyklos*, 59, 17-42.

Ehrlich, M.V. and Overman, H.G. (2020). Place-based policies and spatial disparities across European cities. *Journal of Economic Perspectives*, 34(3), 128–149.

Erdil E. and Çetin, D., (2019). Smart Specialization and R&I Policy Framework in Turkey", in *Smart Specialization Strategies and the Role of Entrepreneurial Universities*, (eds.) N. Caseiro, and D. Santos, 209-33, Hershey: IGI Global.

European Commission (EC) (2009). Knowledge Economists Policy Brief, Smart Specialisation – The Concept. Retrieved October 10, 2017, from https://ec.europa.eu/invest-in-research/pdf/download_en/kfg_policy_brief_no9.pdf

European Commission (EC) (2014). National/Regional Innovation Strategies for Smart Specialisation (RIS3), Cohesion Policy 2014-2020, Policy Document. Retrieved October 10, 2017, from http://ec.europa.eu/regional_policy/sources/docgener/informat/2014/smart_specialisation_en.pdf

European Commission (EC) (2019). *ERA Progress Report 2018*. Accessed 12.11.2021 at https://op.europa.eu/en/web/eu-law-and-publications/publication-detail/-/publication/5641328c-33f8-11e9-8d04-01aa75ed71a1

Farole, T., Rodriguez-Pose, A. and Storper, M. (2011). Cohesion Policy in the European Union: Growth, Geography, Institutions, *JCMS-Journal of Common Market Studies*, 49 (5), 1089-1111.

Fiaschi, D., Lavezzi, A.M. and Parenti, A. (2018). Does EU cohesion policy work? Theory and evidence, *Journal of Regional Science*, 58(2), 386-423.

Fischer, F. and Mandl, H. (2005). Knowledge convergence in computer-supported collaborative learning: The role of external representation tools. *The Journal of the Learning Sciences*, 14(3), 405-441.

Foray, D., David, P. and Hall, B. (2009). Smart Specialisation: The Concept. In: Knowledge for Growth: Prospects for Science, Technology and Innovation. Knowledge for Growth Policy Brief No: 9, EUR 24047. Brussels: European Commission. Retrieved October 10, 2017, from http://ec.europa.eu/invest-in-research/pdf/download_en/kfg_policy_brief_no9.pdf

Foray, D., David, P. A. and Hall, B. H. (2011). Smart specialization from academic idea to political instrument, the surprising career of a concept and the difficulties involved in its implementation. MTEI-Working Paper No. 2011–1.

Foray, D. (2015). *Smart Specialisation: Opportunities and challenges for regional innovation policy*. New York: Routledge.

Foray, D., Goddard, J., Goenaga Beldarrain, X., Landabaso, M., McCann, P., Morgan, K. and Ortega-Argilés, R. (2012). Guide to Research and Innovation Strategies for Smart Specialisations (RIS 3). European Commission, Regional Policy Document. Retrieved October 10, 2017, from http://s3platform.jrc.ec.europa.eu/documents/20182/84453/RIS3+Guide.pdf/fceb8c58-73a9-4863-8107-752aef77e7b4

Forman, E. and Cazden, C.B. (1985). Exploring Vygotskian Perspectives in Education: the cognitive value of peer interaction, in James V. Wertsch (ed.), *Culture, Communication, and Cognition: Vygotskian Perspectives*. Cambridge: Cambridge University Press.

Freeman C., Clark J. and Soete L. (1982). *Unemployment and Technical Innovation: A Study of Long Waves and Economics*, London: Pinter Publisher.

Freeman, C. (1987). *Technology Policy and Economic Performance: Lesson from Japan*, London: Pinter Publisher.

Gilsing, V.A., Nooteboom, B., Vanhaverbekec, W., Duysters, G. and van den Oord, A. (2008). Network embeddedness and the exploration of novel technologies: Technologicaldistance, betweenness centrality and density. *Research Policy*, 37(10), 1717-31.

Graesser, A. C., Person, N.K. and Magliano, J.P., (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, 9, 495-522.

Gremmen, M.C., Dijkstra, J.K., Steglich, C. and Veenstra, R. (2017). First Selection, Then Influence: Developmental Differences in Friendship Dynamics Regarding Academic Achievement, *Development Psychology*, 53(7), 1356-1370.

Gulati, R. (1999). Network Location and Learning: The Influence of Network Resources and Firm Capabilities on Alliance Formation. *Strategic Management Journal*, 20, 397-420.

Hazir, C.S., LeSage, J. and Autant-Bernard, C. (2016). The role of R&D collaboration networks on regional knowledge creation: Evidence from information and communication technologies. *Papers in Regional Science,* 97(3): 549-567.

Hoekman, J., Scherngell, T., Frenken, K. and Tijssen, R. (2013). Acquisition of European research funds and its effect on international scientific collaboration. *Journal of Economic Geography,* 13 (1), 23-52.

Hooijsma, M., Huitsing, G., Kisfalusi, D., Dijkstra, J., Flache, A., and Veenstra, R. (2020). Multidimensional similarity in multiplex networks: Friendships between same- and cross-gender bullies and same- and cross-gender victims. *Network Science, 8*(1), 79-96.

Hutchins, E. (1991). The social organization of distributed cognition. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 283–307). American Psychological Association.

Hutchins, E. (1995). *Cognition in the wild*. Massachusetts: The MIT Press.

Ickes, W. and Gonzales, R. (1996). "Social" cognition and social cognition: From the subjective to the intersubjective. In J. L. Nye & A. M. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 285–308), London:Sage Publications.

Jenson, J. (2010). *Defining and measuring social cohesion*, United Nations Research Institute for Social Development.

Jeong, H. and Chi, M.T.H. (2007). Knowledge convergence and collaborative learning. *Instructional Science*, 35, 287-315.

Kalish, Y. (2020). Stochastic Actor-Oriented Models for the Co-Evolution of Networks and Behavior: An Introduction and Tutorial, *Organizational Research Methods*, 23(3), 511-534.

Lahdelma, T. and Laakso, S. (2016). Network analysis as a method of evaluating enterprise networks in regional development projects, *Evaluation*, 22(4), 435-450.

Lave, J. and Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Leszczenky, L. and Pink, S. (2015). Ethnic segregation of friendship networks in school: Testing a rational-choice argument of differences in ethnic homophily between classroom- and grade-level networks, *Social Networks*, 42, 18-26.

Lundvall, B.A. (1992). *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. London: Pinter Publishers.

Mowery, D.C., Oxley, J.E., and Silverman, B.S. (1998). Technological overlap and interfirm cooperation: implications for the resource-based view of the firm. *Research Policy*, 27, 507-523.

Nelson, R.R. and Winter S.G. (1982). *An Evolutionary Theory of Economic Change*, Cambrdige: Harvard University Press.

Nonaka, I. (1994). *A dynamic theory of organizational knowledge creation*. *Organization Science*, 5(1), 14-37.

Nooteboom, B., Haverbekeb, W.V., Duysters, G., Gilsing, V. and van den Oord (2007). Optimal cognitive distance and absorptive capacity. *Research Policy*, 36, 106-34.

Pellegrini, G., Terribile, F., Tarola, O., Muccigrosso, T. and Busillo, F. (2013). Measuring the effects of European Regional Policy on economic growth: A regression discontinuity approach, *Papers in Regional Science*, 92(1), 217-233.

Resnick, L.B. (1991). Shared cognition: Thinking as social practice. In L.B. Resnick, J.M. Levine and S.D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 1–20). American Psychological Association.

Ripley, R.M., Snijders, T.A.B, Boda, Z., Vörös, A. and Preciado, P. (2020). *Manual for SIENA version 4.0* (version October 19, 2021). Oxford: University of Oxford, Department of Statistics; Nuffield College.

Rogoff, B. (1998). Cognition as a collaborative process. In W. Damon (Ed.), *Handbook of Child Psychology* (pp. 679-744). New York: Wiley.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences*, 2(3), 235-276.

Roschelle J. and Teasley S.D. (1995) The Construction of Shared Knowledge in Collaborative Problem Solving. In: O'Malley C. (eds) *Computer Supported Collaborative Learning.* NATO ASI Series (Series F: Computer and Systems Sciences), vol 128. Berlin: Springer.

Sala-i-Martin, X. (1996). Regional cohesion: evidence and theories of regional growth and convergence, *European Economic Review*, pp. 1325-1352

Schmidt, J. (2019) EU Cohesion Policy: A suitable tool to foster regional innovation? Policy Paper, Bertelsmann Stiftung, Programme Europe's Future. https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/EZ_Policy_Brief_Cohesion_2019_ENG.pdf

Schoenmakers, W. and Duysters, G. (2006). Learning in strategic technology alliances, *Technology Analysis and Strategic Management*, 18, 245-64.

Shin, H. (2017). Friendship Dynamics of Adolescent Aggression, Prosocial Behavior, and Social Status: The Moderating Role of Gender, *Journal of Youth Adolescence*, 46, 2305-2320.

Snijders, T.A.B. (1996). Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology*, 21(1-2), 149-172.

Snijders, T.A.B. (2001). The statistical evaluation of social network dynamics. *Sociological Methodology*, 31, 361-395.

Snijders, T.A.B., van de Bunt, G. G., & Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1), 44-60.

Teasley, S.D., Fischer, F., Dillenbourg, P., Kapur, M. and Chi, M. (2008). Cognitive convergence in collaborative learning, ICLS'08: Proceedings of the 8th international conference on International conference for the learning sciences - Volume 3, pp. 360–67.

Teece, D.J., Pisano, G. and Shuen, A. (1997). Knowledge and the Flow of Information Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18, 509-533.

Van der Wouden, F. and Rigby, D.L. (2019). Co-inventor Networks and Knowledge Production in Specialized and Diversified Cities, *Papers in Regional Science*, 98(4), 1833-1853.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Massachusetts: Harvard University Press.

Wang, C., Hipp, J.R., Butts, C.T., Jose, R. and Lakon, CM (2015). Alcohol Use among Adolescent Youth: The Role of Friendship Networks and Family Factors in Multiple School Studies, PLOS ONE, 10(3).

Wanzenbock, I., Scherngell, T. and Brenner, T. (2014). Embeddedness of regions in European knowledge networks: a comparative analysis of inter-regional R&D collaborations, co-patents and co-publications. *The Annals of Regional Science*, 53(2), 337-368.

Webb, N.M. and Palinscar, A.S. (1996). Group processes in the classroom. In D.C. Berliner & R.C. Calfee (Eds.), *Handbook of Educational Psychology.* New York: Simon & Schuster Macmillan.

WIPO, 2019. T*he Geography of Innovation: Local Hotspots, Global Networks*, Geneva: WIPO.