Of Trees and Monkeys.
The evolution of technological specialization of European regions

M.A. Maggioni[+], E. Marrocu*, T.E. Uberti[+], S. Usai*
[+]CSCC and DISEIS, Università Cattolica del Sacro Cuore, Milano
*CRENoS and SEA, Università degli Studi di Cagliari, Cagliari

> "Think of a product as a tree and firms as monkeys. There are rich and poor parts of the forest. What you want is to have the monkeys jump from the poor part to the rich part. But in some places, the trees are close together, so it is easy for the monkeys to move around. In other places, the trees are far apart: this is where the capabilities that go into making one thing don't help much in making the next thing"
>
> Richard Hausmann, as quoted by Shaw (2010)

Abstract

How do regions develop and evolve along their productive and technological path is a central question in many scientific fields from international economics, to economic geography, from public policy to regional science. Within an evolutionary perspective, we believe that, in general, a given region is most likely to develop new industries or new technologies closer to its pre-existing specialization. Our research builds on an empirical stream of literature, started by Hausmann and Klinger (2007) and Hidalgo et al. (2007), aimed at tracing the world evolution of industrial specialisation, at the country level, following the evolution of export portfolios. We refocus this line of analysis on the regional European technology/knowledge space along the research avenue started by Kogler et al. (2017). We aim at investigating the pattern and the evolution of regional specialisation in the most innovative EU countries in terms of the interaction of three factors: (i) endogenous processes of knowledge recombination and localised technological change, (ii) exogenous technological paradigm shifts and (iii) trans-regional spatial and technological spillovers and networking dynamics.

More specifically, our paper maps the technological trajectories of 198 EU regions over the period 1986-2010 by using data on 121 patent sectors at the NUTS2 level for the 11 most innovative European countries, plus Switzerland and Norway. We map the knowledge space following two distinct and complementary approaches: a micro-level one, based on co-classification information contained in patent documents (as in Engelsman and Van Raan, 1992; Kogler et al., 2017), and a macro-level one, based on conditional co-specialisations of regions in the same patent classes (as in Hidalgo et al., 2007). These two representations of the knowledge space serve as reference bases for understanding the evolution of regional technological specialization, being measured in terms of the sector-region relative technological advantage (RTA), and for modelling its dynamics as a function of spatial, technological and socio-cognitive proximity.

The results show a significant path dependence in the evolution of the regional technological specialisation, whose changes are significantly shaped mostly by phenomena of localised technological change and recombinant innovation. We also find evidence of a significant role played by spillovers and neighbourhood effects in the form of geographic and technological spillovers.

Keywords: Technology/knowledge space, localised technology change, recombinant innovation, European regions, evolutionary economic geography, patent analysis, dynamic spatial models.

JEL codes: O14 O31 O33 O52 R11 R12 C21

1. Introduction

The technological and productive specialization of regions has always been a relevant issue both from a theoretical and empirical viewpoint. While globalisation and the ICT revolution have radically transformed the geography of production – contrary to some early claims about "the Death of Distance" and references to "the World is Flat" (Cairncross, 1997 and Friedman, 2005) and which envisaged the irrelevance of location – they have also spurred the importance of regional specialisation as a relative advantage in an increasingly competitive global arena.

The question about how regions develop and evolve along their productive and technological path has been recently raised in many scientific fields from international economics, to economic geography, from public policy to regional science. Within an evolutionary perspective, we tend to think that a region is most likely to develop new technologies and new industries closer to its pre-existing technological and productive specialization. The analytical framework behind this idea is a blend of two different concepts well established in the theoretical literature on the economics of innovation and technological change, namely (i) the concept of "recombinant growth" developed by Weitzman (1998) and (ii) the concept of "localised technological change" conceived by Atkinson and Stiglitz (1969). According to this framework, while radically new technologies emerge from the recombination of existing technological knowledge, skills and competences, incremental innovations develop along the lines of past technological trajectories by causing local changes in the shape of isoquants rather than global shifts in their position.

The operationalization of these two concepts is not an easy task. Nonetheless, we can nowadays exploit the progress made along the lines of Hausmann and Klinger (2007) who conceived a novel methodology (based on Maximum Spanning Trees) to map the evolution of industrial specialisation at the country level, based on trade flows (see also Hidalgo et al., 2007; Hidalgo and Hausmann, 2008). Other studies (such as Neffke et al., 2013, Rigby, 2013, Boschma, 2014, Kogler et al., 2017) have already applied this methodology at the regional level by using different data, territorial units and geographical setting (country level, EU or USA).

We move along this stream of literature with a specific focus on the European knowledge space (as in Kogler et al., 2017) in order to investigate the evolution of regional specialisation from the middle eighties up to the complete outbreak of the economic crisis. The technological dynamics is, therefore, conceived as the outcome of the complex interaction of two endogenous processes: one of localised technological change (LTC) and the other one of recombinant innovation (RI). At the same time, technological changes may also emerge as a result of exogenous technological paradigm shifts, geographic, technological and socio-cognitive spillovers and network effects. Moreover, we compare and contrast two dimensions/measures of technological interrelatedness, which have been used so far in the empirical literature. The first one, at the macro level, is based on information on co-specialisations in regions, as suggested in the pioneering contribution of Hidalgo et al. (2007); the other one at the micro level, is based on the co-occurrence of patent classes as applied for the first time in Rigby (2013). These two representations of the knowledge space are then used for understanding the evolution of the specialization process, measured in terms of the sector-region relative technological advantage (RTA), and for modelling its dynamics as a function of spatial, technological and socio-cognitive proximity.

Results show that there is a significant path dependence in regional technological specializations, which are shaped mainly by localised technological change and by the exogenous technological shift of the European technological frontier. On the contrary the phenomenon of recombinant innovation is less decisive unless it is interacted with cases of strong specialisation. We also find evidence of spillovers induced by both geographic and technological connectivity.

The paper is organised as follows. The next section connects this analysis to the established literature in the field, section three describes the methodology used to derive Maximum Spanning Trees from technological interrelatedness matrices, and presents some stylised facts about the European technological space and its regional evolution over time. The fourth section introduces the main explanatory variables, while the fifth section presents the estimation strategy and the empirical analysis results. Concluding remarks and future research agenda are in the final section.

2. Literature review

2.1 Analytical background

The prevalent model of technological change used in empirical analysis – the knowledge production function (Griliches, 1979) – assumes that the greatest source generating knowledge, besides human capital and skilled labour, is public and private R&D. This empirical model has been applied at different dimensions of economic systems: from the micro level of firms and plants to the macro level of sectors, regions and nations. According to this view, new ideas and knowledge, and their plastic conversion into an orthodox production function, are the simple output of the interactions of some scientific and technological research effort.

This mechanical idea of knowledge creation is, however, not completely satisfactory and even Griliches himself in his conclusion acknowledges that " We need more research on … how to conceptualize and estimate technological distance between firms and industries and the associated notions of externalities and spillovers in research" (Griliches, 1979, p. 43).

A more convincing view is that knowledge production is the result of research efforts along an evolutionary process (Boschma and Frenken, 2011) which can be due to a blend of localised technological change (Atkinson and Stiglitz, 1969) and recombinant innovation (Weitzman, 1998). Along this complex process of knowledge formation and diffusion, geography may play a decisive role because of the nature of local public good of ideas, that is the presence of knowledge spatial stickiness, (as explained in the pioneering survey by Audretsch and Feldman, 2004, and more recently by Feldman and Kogler, 2010). Moreover, technological relatedness among products and innovations can affect the nature and scope of local knowledge spillovers within a region. In other words, new regional competences and technologies depend on pre-existing scientific and technical knowledge, skills and practical experiences. This is because firms in different but related technologies are more able to gain (that is, to profit) from spillovers, than firms in unrelated activities. Secondly, the emergence of new technologies from existing and related technologies depends on the current sectoral composition of a regional economy since it provides the basis for the recombination of existing ideas for starting new knowledge paths and, potentially, a process of structural change.

Evidence of the importance of industrial history or regions in conditioning their future specialisation portfolio has been provided in several case-studies (such as Bathelt and Boggs, 2003; Glaeser, 2005; Boschma and Wenting, 2007, Colombelli et al., 2014, Feldman et al., 2015). More interestingly from our point of view, Hausmann and Klinger (2007) and Hidalgo et al. (2007) propose an original methodology in order to identify technological relatedness thanks to export patterns and composition at the country level.

Hidalgo and Housman use a persuasive metaphor of their methodology: products are trees and forests compose the economic structure of countries; firms are, instead, monkeys that live on different trees and exploit those products. Growth dynamics can be, thus, described as the movement of firms from a poorer part of the forest to rich parts of the forest, where trees have better fruits and develop faster. This metaphor is essential to appreciate the concept of interrelatedness: "if this forest is heterogeneous, with some dense areas and other more-deserted ones, and if monkeys can jump only limited distances, then monkeys may be unable to move through the forest." (Shaw, 2010).

Consequently, the composition and the relative density of a forest, that is the economic structure of a country/region, is crucial in determining the orientation and the pace of development of countries/regions in the short and the long run. Hidalgo et al. (2007) employ this method in order to show that rich countries specialize in more densely connected parts of the product space, whilst poor countries develop mainly products in the more peripheral and isolated areas of the same space. As a result, rich countries have more opportunities to sustain economic growth in the long run, thanks to a fruitful process of structural change.

This process is analytically described in a regional evolutionary framework by Boschma and Frenken (2011). They introduce the concept of regional branching, to identify those cases when a new variety is rooted in related activities in a region. Regional branching may occur either because an innovation grows out of an old technology, or because a new idea results by the recombination of competences and experiences of different technologies (thus being perfectly compatible with our two main explanatory factors: namely recombinant innovation and localized technological change).

Relatedness and proximity can be crucial in favouring changes, not only in the geographic and the technological space, but in other dimensions too. Boschma (2005) presents a taxonomy which includes social, organisational and institutional proximities as other potentially favouring factors for knowledge spillovers.

2.2 Productive and technological specialization at national/regional level

Hidalgo et al. (2007), as mentioned above, use export data in order to analyse the productive specialisation pattern of countries and focus on the shifts of country's export portfolio along time as a proxy of structural change and industrial dynamics. Export datasets have the great advantage of being finely disaggregated and available for long time series. However, there is another database with similar characteristics, the patent statistics. Exactly for this reason, patents have been diffusedly employed to analyse national technological specialisation and its changes as a proxy of industrial structure and dynamics (Pavitt, 1988; Archibugi and Pianta, 1992). In particular international cross-patenting has been frequently used as a national technological indicator and as an indirect measure of a country's productive specialization (Soete and Wyatt, 1983; Leoncini et al., 1996; Paci et al., 1997).

More recently the focus has shifted from countries to regions, since technological knowledge may have a tacit nature and consequently can be strongly associated with local capabilities, institutional setting and social capital. Regions may, therefore, accumulate specific competences and intangible assets, which provide spatially and cognitively bounded learning opportunities for local firms (Lawson, 1999; Breschi, 2000, Greunz, 2003; Moreno et al., 2005).

This research avenue has recently regained momentum thanks to some original contributions, which have adapted at the regional level the methodology proposed by Hidalgo et al. (2007). Some studies have primarily focused on the impact of technological relatedness on the opportunities to grow by provinces in Italy (Boschma and Iammarino, 2009) and in Spain (Boschma et al., 2012).

Other works are more directly oriented to the issue of industrial branching within regions and how this is influenced by the structure, or the relatedness, of local economic environment. Neffke et al. (2011) study products entry and exit in 70 regions in Sweden by looking at employment data and measuring technological relatedness thanks to an original dataset on product co-occurrences in plants.

Boschma et al. (2014) investigate, thanks to USPTO patents, the role of technological relatedness in pushing and orienteering technological change in 366 US cities (MSA) from 1981 to 2010. They find that the presence of technological relatedness may play a crucial role by increasing the entry probability of a new technology and decreasing the exit probability of an existing technology. They use two different methods to build the relatedness indicator. The main method follows the product space framework proposed by Hidalgo et al. (2007), where two technologies are considered related if they have a revealed technological advantage within the same US city. The second method, used to test the robustness of the results, is based on Hall et al. (2001) patent classification and a normalized co-occurrence analysis.

A similar method is at the heart of the contribution by Rigby (2013) who studies the evolution of knowledge space in the same sample of 366 MSA from 1975 to 2005. Technological relatedness is constructed by using patent citations, it is given by the probability that a patent in class *j* will cite a patent in class *i*,. Such probabilities are computed on the basis of the links within the knowledge space. The analysis shows that average relatedness between US patents in thirty years has almost doubled since patents are increasingly concentrating in fewer technology classes, which are becoming more proximate (or related). As far as the determinant of entries and exits of cities from patent classes, the expansion of the knowledge core, depends crucially upon the proximity of new technological possibilities to the set of existing specializations. Most interestingly, estimations show that other dimensions of proximity, other than technological one, play a role: diversification is influenced also by the knowledge available in socially closer locations, where social proximity is measured in terms of co-inventors links.

US Metropolitan areas are also at the centre of Essletzbichler (2015) analysis, even though the relatedness measure is based on input-output linkages between industries rather than patent or products co-occurences. Nonetheless, results confirm that technological relatedness is positively related to previous industry portfolio membership and industry entry and negatively related to industry exit.

Finally, Kogler et al. (2017) present the latest contribution in this quite rich recent avenue of research. Their contribution opens the analysis to the European context and most importantly across countries. They

use patent co-classification data to measure the proximity between all pairs of International Patent Classification categories in order to map and track the evolution of knowledge space from 1981 to 2005 in 213 NUTS2 regions of EU15. They find that, as in the US, in Europe knowledge specialization has increased significantly along time. They also show that entry, exit and selection processes over space and time are influenced not only by the proximity to the knowledge core of the region but also to knowledge spillovers from neighboring regions.

Another stream of literature discuss the issues whether knowledge is diffused and exchanged either through a diffusive pattern based on spatial contiguity, or according to intentional relations based on a-spatial networks (Maggioni et al., 2007; Miguélez and Moreno, 2017).

According to the first pattern, the geographic selection process leading to a hierarchical structure of the location of innovative activities goes together with an increasing role of 'unintended' spatial knowledge spillovers within the network of firms, universities and research centres, all located in neighbourhood areas. Thus, relevant regions present both an 'attractivity' potential and a 'diffusive capacity' (Acs et al., 2002). The exchange of knowledge among firms is facilitated by their geographic proximity, given that knowledge has, in part, a tacit nature that tends to bind the spatial scope of spillovers (Jaffe et al., 1993; Audretsch and Feldman, 1996).

According to the second pattern, knowledge is mainly exchanged according to a voluntary 'barter' and increased through learning by interacting procedures, within specialised networks which are intentionally established between crucial nodes (Cowan and Jonard, 2004). This approach has highlighted that interfirm exchanges can also be mediated by other dimensions of closeness, which may have an a-spatial nature, such as cognitive, institutional, or organizational proximity (Torre and Gilly, 2000; Boschma, 2005) and, more relevantly, that interactions among economic agents create social links that, over time, tend to evolve into wider networks, which are likely to facilitate the future exchanges of knowledge and moderate the adverse effects of other distances (Boschma and Frenken, 2010). A growing body of empirical research has extensively analyzed the characteristics of networks that are expected to prompt innovation diffusion by considering various forms of connections among agents. These include participation in research programs (Autant-Bernard et al., 2007; Maggioni et al., 2007; Balland, 2012), co-patenting (Cantner and Meder, 2007; Maggioni et al., 2007; Cassi and Plunket, 2013), citations (Maurseth and Verspagen 2002; Paci and Usai 2009), co-publications (Ponds et al., 2007), applicant-inventor relationships (Maggioni et al., 2011; Picci, 2010), and human capital mobility (Miguelez and Moreno, 2013; Breschi and Lissoni, 2009).

This paper try to put together all these streams of literature and apply an encompassing framework to the analysis of the evolution of the technological specialization of European regions.

3. The European knowledge space

In this study we depict the most salient traits of the European knowledge space featured over the period 1986-2010 by means of the maximum spanning trees (MST) method. The MST are constructed following two different approaches: a macro *top-down* approach based on co-specialization (named HH, after Hidalgo et al., 2007) and a micro *bottom-down* approach based on co-classification (named CC). Hence, for each region and each time period we adopted the same MST (in either the HH or CC version), since these are defining the same knowledge and technological interrelatedness at the European level.

MST are built by using data on the number of patent applications filed at the European Patent Office (EPO) classified by priority year and by inventor's region for 198 NUTS2 regions (r = 1, …, R=198) in Europe (EU13), belonging to the most innovative countries in Europe, recording 97% of total European patents in the period 1986-2010 (see table A1 in the Appendix).

Since patenting activity at the regional level is quite irregular over time, we smooth the patent variable by computing five-year period averages. Thus, our analysis is articulated in 5 (t = 1, …, T=5) five-years periods[1].

In order to deal with the sectoral dimension of technological interrelatedness, we focus on 121 International Patent Classification (IPC) classes at the second hierarchy level (i = 1, …, I=121) and 8 IPC sectors (z = 1,… Z=8). Namely, A. Human necessities; B. Performing operations, transporting; C. Chemistry, metallurgy; D.

[1] Time time intervals are as follows: 1986-90 (T1), 1991-95 (T2), 1996-2000 (T3), 2001-05 (T4), 2006-10 (T5).

Textiles, paper; E. Fixed constructions; F. Mechanical engineering, lighting, heating, weapons, blasting; G. Physics; H. Electricity.

3.1 MST: HH *vs* CC

In order to describe the technological interrelatedness between IPC classes for EU13 regions, we started from the "innovation space" , a notion derived from the "product space" as defined by Hidalgo et al. (2007). The "innovation space" is, in principle, a connectivity matrix which shows how closely interrelated is an IPC class with another one.

In our case we select two ways for measuring such a connectivity. This is why we define 2 sets of connectivity measures, each consisting of a 121x121 matrix whose rows and columns represent each IPC classes and each off-diagonal cell represents two distinct ways to measure the technological connectivity between a given pair of IPC classes, alternatively based on the macro *top-down* or the micro *bottom-up* perspective.

Since a matrix can easily be interpreted as a network, in which each IPC class is a node and the connectivity measure between a couplet of nodes is a link, in the following paragraphs we will use these terms as synonyms.

In particular, the macro top-down approach (HH) is implemented using a similar method to Hidalgo et al. (2007), which relies on conditional probabilities of a region being specialized in IPC class *i* given that the region is also specialized in in IPC class *j*. Regional specialization in a given IPC class is measured in terms of Revealed Technological Advantage (RTA). This is computed as the proportion of a region's patents by inventors (*pat*) in a given IPC class *i* in period *t* divided by the proportion of European patents in the same IPC class in the same period. RTA is thus a measure of relative specialization of a given region in a specific IPC class. Formally:

$$RTA_{irt} = \frac{\frac{pat_{ir}^{t}}{pat_{ir}^{t}}}{\frac{pat_{iEU13}^{t}}{pat_{iEU13}^{t}}}$$

Given the conditional probabilities $P(RTA_i|RTA_j)$ and $P(RTA_j|RTA_i)$ of a region being specialized in IPC class *i* given that the region is also specialized in in IPC class *j*, each element of the **HH** connectivity matrix at time *t* hh_{ij} is equal to:

$$hh_{ij}^{t} = min\{P(RTA_i^t|RTA_j^t), P(RTA_j^t|RTA_i^t)\}$$

In this definition, as in Hidalgo et al. (2007), we consider the minimum between the two conditional probabilities because we need to weigh the strength of the connectivity between two specific IPC classes by taking into account the fact that one IPC class may be more diffused than the other one.

The micro bottom-up approach (CC) is implemented using a similar method to Engelsman and Van Raan (1992). According to this approach, "two technology classes are considered to be technologically related if they occur frequently together as technology classification codes on the same patent" (p. 6). Each element of the **CC** connectivity matrix at time *t* cc_{ij} is computed for each period as the number of EU13 patents in which a given couplet of IPC classes *i,j* is jointly occurring. Analytically:

$$cc_{ij}^{t} = \sum_{pat=1}^{N} pat_{i,j}^{t}$$

Since both **HH** and **CC** matrices are very dense and links values are very heterogeneous[2], we decided to focus our analysis only on key technological relations underlying the European technological space. In order to do so, following Hidalgo et al. (2007), for each interval of time we identified a European MST, whose nodes are 121 IPC classes and links include exclusively the most relevant technological interrelation between a couple of IPC classes.

The procedure to create the MST (both for the macro HH and the micro CC version) is iterative and starts with the identification of the maximum link value in each connectivity network.[3] Once the maximum value has been selected, we establish a link between that couplet of IPC classes, or nodes. Secondly, by focusing on the identified dyad, we search for a further node to be connected to that dyad in order to form a triad. The link is identified by searching for the maximum value of all links attached to one of the two nodes of the dyad. The procedure is iteratively repeated by adding nodes (i.e. IPC classes) and links, until all IPC classes are included in the MST (which, by definition is a minimally dense network of *N* nodes and *N-1* links). The final MST structure for initial (T1: 1986-1990) and the last (T5: 2005-2010) time periods considered, under both HH and CC perspectives, are depicted in Figures 1 and Figure 2.

In order to detail the possible relatedness among technological competencies, for each MST we computed a "density" kind of measure for each one-level digit (*z*=1, 2, …8), i.e. the ratio of the existing links within each one-level IPC sector and all potential links that would characterize an ideal MST of similar dimension:

$$density_{z,MST}^{t} = \frac{L_z}{(N_z - 1) \times 2}$$

The density average values are reported in Figure 3. On average, over time, only IPC sector H (i.e. Electricity) is above mean in both HH and CC MST indicating that more than a half of links actually existing are also those links that will cover a MST with the same dimension (i.e. fine nodes network). For sector H knowledge appears to be relatively concentrated showing a low level of knowledge externalities.

On the contrary, IPC sectors A (i.e. Human Necessities) and C (i.e. Chemistry and metallurgy) show a similar pattern in both MST. These ratios seem to show related variety patterns, since the existing links are much lower than a potential MST connecting exclusively these sectors. These results could also point out a limit in IPC sectors definition, whose distribution of second hierarchy level is very uneven. The remaining sectors, on average, show different patterns in all MST (Figure 3).

The evolution of the densities metrics in each MST highlights interesting patterns. IPC sector B (Performing operations, transporting) shows almost always unrelated connections in the HH MST, and exactly the opposite happens in the CC MST (see Figure 3 for average values). Sector D (Textiles, paper) shows the opposite pattern, i.e. very related in the HH MST and very unrelated in the CC MST.

Since we are interested on the role played by each region on the European technological space, in the empirical analysis we keep the MST (both CC and HH) as two complementary representations of the technological interrelatedness at the European level, while the specific values of each regional specialization, in terms of RTA in a given IPC class are region-specific.

3.2 HH and CC: a correlation analysis

By looking at the temporal evolution of HH and CC MSTs from 1986 to 2010 it is evident that significant changes in the MST structure have occurred. Thus, looking for a standardized measure of

[2] It is worth noting that strictly positive values of connectivity are increasing over time. In particular, in the HH connectivity matrix, positive values are 98% (or more) of all possible values. On the contrary, in the CC connectivity matrices, positive values are never more than 82% of the possible values, and the trend is non linear. Initially only 24% of the possible links are present; in the second and third periods percentages nearly double (i.e. 40% and 47%), while in the fourth and fifth periods percentages increased tremendously (i.e. 70% and 82%). Finally, there was a drop in the last period (i.e. 58%, but this could be related to the structural slowdown in patents applications in the most recent years). Since some matrices are very dense (such as in the HH connectivity matrix) and some matrices are very variable over time (such as in the CC connectivity matrix), we decide to perform the empirical analysis focusing on the key and synthetic frontier (the one based on the MST structure) of the European knowledge-technology space.

[3] A simple, intuitive description of the procedure adopted to build the MST is sketched in the Appendix.

correlations between different networks we resolved in using the Quadratic Assignment Procedure (QAP), to calculate the extent to which the pattern of links in one network is similar to the pattern in another network. Standard correlation is not appropriate for dyadic data because this type of data are not independent of each other. QAP controls for the non-independence of the cases through the use of several random permutations of rows and columns of the original matrix though a Monte Carlo procedure, thus it allows to rule out spurious correlations (Krackhardt, 1988).

Table 1 shows some interesting results: all QAP correlation coefficients are significant, with the HH MST ones being smaller than the ones computed for the CC MST. Both HH and CC coefficients indicate positive autocorrelation, whose strength tend to decrease over time (i.e. the correlation between the MST_{HH}^{1} with the MST_{HH}^{2} is larger than the value of the correlation between the MST_{HH}^{1} with the MST_{HH}^{5}).

The reduction in the correlation values may be interpreted as a sign of the incremental nature of technological change, as time passes the technological frontiers keeps modifying on the basis of the previous one. However, by comparing the correlation coefficients of different time-contiguous couplets (i.e the correlation between the MST_{HH}^{1} and MST_{HH}^{2} with the correlation between MST_{HH}^{2} with the MST_{HH}^{3} etc.) it is not possible to detect a clear pattern of either an increase or a reduction in the speed of technological change.

Finally, it is interesting to compare HH and CC MST: simultaneous MST are positively but weakly correlated, following a non-linear trend over time. Positions of nodes in CC and HH MST are pretty dissimilar, i.e. IPC classes are displaced differently in the MST, and the effects on the technological specialization of regions should be different. It seems that the technological interrelatedness structures at the European level, when measured from a macro and a micro perspective, are positively related but dissimilar. In the empirical analysis we use both the HH and the CC MST in order to assess to what extent results are sensitive to the use of the two different approaches – top-down and bottom-up – adopted to describe the European knowledge space.

4. Description of variables and proximity measures

As stated in the introduction, the main aim of this study is to explain the evolution of the technological specialization of the European regions as a function of localized technological change (LTC), recombinant innovation (RI), exogenous technological shift (ETS), while accounting for connectivity factors and persistency over time. Regional technological specialization is measured in terms of RTA, as described in the previous section. Formally:

$$RTA_{irt} = f\,(LTC_{irt}, RI_{irt}, ETS_{irt}, \textit{proximity factors}, \textit{persistency})$$

where *i* indicate the 121 IPC patent classes, *r* the 198 regions and *t* the 5 five-year time periods.[4]

4.1 The explanatory variables

All explanatory variables are computed on the basis of the MST for both the HH and the CC approaches. In what follows we provide a detailed description of the procedure followed to construct each variable.

Localized Technological Change

LTC is expected to provide empirical support to Atkinson and Stiglitz (1969) claim that "the different points on the [production possibilities] curve represent different processes of production, and associated with each of these processes there will be certain technical knowledge specific to that technique. Indeed, both supporters and critics of the neoclassical theory seem to have missed one of the most important points of the activity analysis (Mrs. Robinson's blueprint) approach: that if one brings about a technological improvement in one of the blueprints this may have little or no effect on the other blueprints. If the effect of technological

[4] A complete description of all variables, along with basic descriptive statistics, is reported in Table A2 in the Appendix.

advance is to improve one technique of production but not other techniques of producing the same product, then the resulting change in the production function is represented by an outward movement at one point and not a general shift. … In reality we should expect that a given technical advance would give rise to some spillovers and that several techniques would be affected" (Atkinson and Stiglitz, 1969, p. 573).

In order to translate the direct "effects of technological advance" from a technology to a contiguous one, we exploited neighbourhood concepts drawn from network analysis and graph theory.

Given a set of nodes $\mathcal{N} = \{n_1, n_2, \dots, n_{121}\}$, there are several paths, with different lengths, connecting a given pair of nodes. The shortest path between two nodes *i* and *j* is named geodesic distance and is denoted as g_{ij}. If $g_{ij} = 1$, nodes are adjacent, indicating that there exists a direct link between them, otherwise if $g_{ij} > 1$, nodes are not directly linked and the number indicates the smallest length connecting them. Hence, to detect the direct, or local, technological effects we used the concept of adjacency of nodes. In this case, the local neighbourhood is defined as $g_{ij} = 1$. Therefore, LTC is obtained by computing per each IPC class, each region and time the summation of RTA of nodes directly adjacent on the MSTs (in both the HH and the CC version). Formally:

$$LTC_{irt} = \sum_{j=1}^{N-1} RTA_{jrt} | (g_{ij} = 1, MST_t)$$

In order to take into account possible effects on the regional specialization in a given IPC class arising from indirect technological connectivity, we also constructed the variable *NEighbouring ALbedo (NEAL)*. It is built on the same principles of LTC. In fact, as we did for LTC variable, we exploited the concepts of geodesic distances higher than 1 in order to evaluate the indirect effects of technological advance.

For each region, sector and time we computed NEAL as the sum of the RTA values of second and higher order neighboring classes and weighting them for each geodesic distance from the target IPC class (i.e. node *i*) to any other IPC in the MST. Formally:

$$NEAL_{irt} = \sum_{j=1}^{N-3} \frac{RTA_{jrt}}{g_{ij}} | (g_{ij} > 1, MST_t)$$

Recombinant Innovation

RI is expected to provide empirical support to Weitzman (1998) claims that "[In the knowledge production function approach] 'New ideas' are simply taken to be some exogenously determined function of 'research effort' in the spirit of a humdrum conventional relationship between inputs and outputs. Essentially, this approach represents a theory of knowledge production that tries to do an end run around describing the creative act that produces the new ideas. If new ideas are postulated to be a function of something—for example, research effort—then what is the nature of the functional relationship? Is production of knowledge a process that can be modelled by analogy with fishing new ponds or discovering new oil reserves? It seems to me that something fundamentally different is involved here. When research effort is applied, new ideas arise out of existing ideas in some kind of cumulative interactive process that intuitively seems somewhat different from prospecting for petroleum. To me, the research process has at its centre a sort of pattern-fitting or combinatoric feel. The core of the analytical structure is a theory of innovation based on analogy with the development of new cultivated varieties by an agricultural research station. 'Recombinant innovation' refers to the way that old ideas can be reconfigured in new ways to make new ideas" (Weitzman, 1998, p. 332-333).

In order to reproduce the concept of "new recombination of old knowledge", from social network analysis we adopted the concept of betweenness, i.e. an analytical measure of the strategic role played by nodes lying between geodesic paths connecting other nodes (Freeman, 1979).

To compute the RI variable we adopted a modified version of betweenness centrality (Freeman, 1979) for each IPC class (i.e. a node in a MST) in each region and period according to a three-step procedure.

Firstly, for each region in any period within the network $\mathcal{N}$ we distinguish those nodes exhibiting a value of $RTA_{irt} \geq 1$ from those having a $RTA_{irt} < 1$.

Secondly, we compute the number of times a node *i* is lying on the geodesic paths linking nodes *j* and *k* whose RTA_{irt} is higher or equal to 1.

Finally we weighted each value by a constant value, $(N-1)\times(N-2)\times 2$ (i.e. 28560) in order to normalize each value for the European MST. Analytically:

$$RI_{irt} = \sum_{j\neq j}^{N-1} \frac{g_{jk}(n_i)}{(N-1)\times(N-2)\times 2} |RTA_{jrt}, RTA_{krt} > 1, MST_t)$$

Exogenous Technological Shift

All previous variables are computed considering the structure of the MST and RTA with reference to the same time period. However, in this way we are unable to disentangle, for each region and IPC classes, the effects of its previous technological structure from those arising from exogenous changes in the European "technological frontier" which, by definition is an exogenous phenomenon from a regional viewpoint.

For this reason, we computed ETS as a variable similar to LTC, but with a relevant difference: while in the LTC variable the RTA and the MST are contemporaneous, in computing the ETS variable the MST is one period ahead (t+1) with respect to the value of the RTA(t). In this way we are able to see whether past specialization of previously distant IPC sectors, which only recently have become technologically proximate, played a role in determining the relative specialization of a specific IPC class in a given region. Being used in conjunction with LTC, which controls for the effects driven by past MST, this variable is able to show the effect of exogenous technological change in shaping a region's technological specialization. Formally:

$$ETS_{irt} = \sum_{j=1}^{N-1} RTA_{jrt} | (g_{ij} = 1, MST_{t+1})$$

Proximity factors

In this study we consider proximity along both the geographic and the technological dimensions.

The geographic matrix (W_G) is computed as the inverse of the distance matrix between centroids of each region in the sample. The technological matrix (W_T) is computed on the basis of socio-cognitive data. Each element of the W_T matrix measures co-inventorships for couplets of regions. Differently from the geographic matrix, it changes over time. Both matrices have been normalized by means of the maximum eigenvalue normalization (Kelejian and Prucha, 2010).

5. Empirical analysis

5.1 The econometric model

The econometric analysis is performed by estimating the following baseline specification for the regional technological specialization model:

$$RTA_{irt} = \rho RTA_{ir,t-1} + X_{ir,t-1}\beta + W_G X_{ir,t-1}\gamma + W_T X_{ir,t-1}\delta + \tau_t + C_{ir} + e_{irt}$$

The specification above is a dynamic SLX (Spatial Lag of X) model, which allows us to account for both time, geographic and technological dependence. The SLX model was proposed by Elhorst (2014) and Vega and Elhorst (2015) as a very flexible approach to account for the existence of spatial spillovers. In the regression above we account for time dependence by including the lagged dependent variable at time *t-1*.

The matrix $\boldsymbol{X}$ includes the main explanatory variables (LTC, NEAL, RI and ETS), while the matrices $\boldsymbol{W_G X}$ and $\boldsymbol{W_T X}$ include spatial and technological lags of the explanatory variables, computed by pre-multiplying the explanatory variable matrix $\boldsymbol{X}$ by the geographic and the technological matrix, respectively.

Such terms are expected to account for regional dependence induced by the existence of knowledge spillovers. As emphasized by LeSage and Pace (2009), overlooking such kinds of spillovers may result in biased and inconsistent estimators. Moreover, their existence posits unavoidable challenges to regional policy-makers, as we discuss in the next section. Within the spatial specification reported above, the effect of a given variable becomes more complex: its total effect can be decomposed into a *direct* component, due to changes occurred in a region's own variable, and an *indirect* or spillover one, caused by changes in the same variable taking place in neighbouring regions. In the specification above the parameter vector β includes the short-run direct effects, while the parameter vectors γ and δ include the short-run indirect effects due to the geographic and the technological proximity. It is worth noting that in the SLX specification spillovers are local in nature. Moreover, differently from other widely applied spatial specification (such as the spatial autoregressive one) in the SLX model the ratio between the direct and the indirect effect is not constrained to be the same across the explanatory variables.

In order to attenuate the problem of endogeneity, which could arise because of possible simultaneity, all the explanatory variables are included in the model with a one-period lag. Given that such lag refer to the average over the previous five years, it is supposed to be sufficiently long to break the correlation between the error term and each of the regressors.

The model includes also time period fixed effects (τ_t) to account for macroeconomic shocks common to all region in the sample and country fixed effects (C_{ir}), which are supposed to take into account both observed and unobservable institutional factors.

5.2 Results

The main results of the econometric analysis are reported in Table 2 and Table 3. In the first table we present the estimation of three different specifications, while in table 3 we report the long-run direct and indirect effects that each explanatory variable exerts on the regional technological specialization. Each specification is estimated for both the HH and CC approaches. This would allow us, in principle, to assess whether and to what extent the estimated effects depend on the different representation of the knowledge space.

Model (1) in Table 2 reports a basic version of the econometric model presented in section 5.1, as in this model we do not include the terms accounting for geographic and technological proximity. Under both the HH and the CC scenarios, results point out that the technological specialization of a region in a given IPC class is significantly influenced by its past specialization, thus confirming the path-dependent nature of scientific and technological discovery.

The relevance of localized technological change – i.e. the specialization in contiguous IPC patent classes, the Neighbouring Albedo – which measure the indirect neighbours' technological specialization – and the exogenous technological shift of the European technological frontier exhibits positive and significant coefficients under the HH scenario. Similar results are found for the CC scenario, with the exception of the Neighbouring Albedo variable, which turns out to be not significant.

It is worth noting that to account for exogenous technological shift we have also included an additional variable accounting for shifts occurring in second or higher order (more distant) technological neighbours. The coefficient of such additional term is not significant, indicating that regional specialization is driven by exogenous technological shifts occurring in the most related sectors.

Focusing on Recombinant Innovation, results point out that it is does not seem to be a relevant determinant of the regional technological specialization in the case of the model based on the HH approach, while it exhibits an unexpected negative coefficient in the case of the CC one. Such a difference might be interpreted as a measure of the stronger impediments (when compared to an economic system as a whole) that, at the micro level, an individual inventor or entrepreneur) has to face in recombining previously known but distant pieces of knowledge in order to produce a successful innovation.

However, it could be also reasonable to hypothesize that the cumulative interactive process of recombining ideas features a relevant nonlinear behaviour so that, in order to be effective in enhancing regional technological specialization, a region has to have already acquired a certain level of comparative advantage in a given IPC class. To test this hypothesis we re-estimate the model by including an interactive term between RI ad a dummy variable which takes value equal to 1 if RTA>1 and zero otherwise. Results obtained by estimating model (2) confirm our hypothesis, as the positive interactive term's coefficient

indicate that a stronger specialization in a given IPC class is necessary to take advantage of the possible recombination of technologically distant specializations. With respect to model (1), the results for the other coefficients remain mainly unchanged, with only the ETS coefficient being now significant at a level which marginally exceeds the 10%.

Model (3) extends the previous model by including the proximity terms, which are supposed to account for the existence of geographic and technological spillovers arising from interregional flows of scientific and technological knowledge. In this specification, the estimated coefficients for the non geographic or technological lagged variables remain mainly unchanged with respect to model (2), indicating that they are robust to the inclusion of the connectivity covariates.

Focusing on the spatial proximity terms, the coefficients associated with LTC and RI variables display negative values, under both the HH and the CC scenario. This result can be interpreted as evidence of geographic competition (or locational shadowing) phenomenon, as described by Arthur (1990) and Maggioni (2002). In other words, the relative specialization of geographically proximate regions in a given IPC class deters the specialization of the region under analysis in the same class since inventors may find better alternative locations nearby. On the contrary, the existence of highly specialized regions (RTA > 1) in the neighbourhood yield beneficial effects by the recombining process of existing knowledge.

A different story is told when proximity is measured along the technological dimension by the extent of co-patenting in a given IPC class. In this case the coefficient of the technologically lagged LTC variable is positive both in the HH and in the CC scenario, while negative or not significant coefficients are estimated for the technology lag of both NEAL and RI variables. The technological specialization of a region in a given IPC class thus seem to be favoured by the specialization in directly connected technologies of the main technological regional partners but this is not the case when a more general measure of technological interrelatedness (NEAL) which measures the ability to combine distant technology is taken into account. Again, here we may find evidence of a technological competitive dynamics between technologically interacting regions.

Table 3 reports the long-run direct and indirect effects for the main explanatory variables. The effects are estimated on the basis of the equilibrium version of model (3) in Table 2

The results indicate that, in terms of direct effects, localized technological change is the most effective variable in enhancing regional technological specialization according to both the HH and CC representations of the knowledge space. Because of spatial competition, the own region's efforts are considerably reduced, as evidenced by the negative spatial indirect effect. On the other hand, technological proximity exhibits larger counter balancing effects, especially in the case of the HH scenario. Overall, the total LTC effect is quite sizeable under the HH scenario (0.13), while it is much more contained (0.0018) under the CC one. In order to assess the relative magnitude of such effects, it worth noting that the sample average of the RTA variable is 1.08 (median 0.45). A generalized increase of one standard deviation in the LTC variable would result in an increase of 0.53 in RTA under the HH scenario and 0.008 under the CC one.

The NEAL variable, although does not exhibit sizeable total effects (-0.0009 for HH, -0.0049 for CC), seem to exert a detrimental effect on regional technological specialization, which originates from the negative indirect effect due to technological proximity. This result indicate the existence of competition among regions along the socio-cognitive dimension.

Recombinant Innovation, under both the HH and CC scenarios, does not seem to favour technological specialization at the local level (total effect: -0.0055 for HH, -0.0031 for CC). However, when focusing on highly specialized territories, those with RTA>1, the effects turns out to be positive, 0.0089 (HH) and 0.0075 (CC). In this case, a region's own efforts in the creation of new knowledge by means of recombining pieces of the existing one are amplified by the positive role played by spatial proximity. Being neighboured by highly specialized regions active in ideas-recombining processes favours innovation and enhances technological specialization.

The evidence provided on the existence of significant differences in the role played by regional internal and external determinants of technological specialization posits relevant challenges to policy-makers as a region's innovative efforts could be overturn or seriously weakened by competition processes. This point to the need for coordination policies aimed at fostering regional collaboration along both the spatial and technological-scientific dimensions.

Overall, our analysis has provided stimulating insights on the determinants of technological specialization of the European regions and on how their effects change over time and when taking into proper account the role of geographic and technological connectivity among regions. The analysis has also shown that the effects could be different according to the way in which the knowledge space is represented,

with the top-down macro approach assigning a prominent role in the long run to localized technological change.

The results provided so far could be made more robust by tackling the limits of the current analysis. In future research, our intention is to account for the pronounced asymmetry displayed by the RTA variable and to perform a subsample analysis to deepen our understanding on the evolution of technological specialization according to different characteristics of the regions and of technological macro-classes.

6. Conclusions

The technological and productive specialization of regions has always been a relevant issue both from a theoretical and empirical viewpoint. It has been analysed in many scientific fields from international economics, to economic geography, from industrial economics to regional science.

In this paper we have presented an empirical analysis focusing our attention on 198 NUTS2 regions, belonging to the most 11 innovative EU countries plus Switzerland and Norway. Data on patents classified in 121 IPC sectors observed over the period 1986-2010 have been used to map the knowledge space according to two, possibly complementary, approaches: a micro level one, based on co-classification information contained in patent documents (Engelsman and Van Raan, 1992; Kogler et al., 2017), and a macro level, based on conditional co-specialisations of regions in the same IPC Classes (Hidalgo et al., 2007).

On the basis of these two representations of the knowledge space we have investigated the evolution of the specialization process, measured in terms of the sector-region revealed technological advantage The analysis has been carried out within the theoretical framework based on “recombinant innovation” (Weitzman, 1998) and “localised technological change” (Atkinson and Stiglitz, 1969), which argues that new technologies emerge from the recombination of existing knowledge, skills and competences, and that technological, spatial and social proximities are crucial to develop new knowledge.

Our empirical analysis, carried out by estimating dynamic spatial models, has provided convincing evidence on the role played by localised technological change, knowledge recombination, exogenous technological shifts, and spillovers arising from both geographic and technological regional proximity. As such spillovers can also have adverse effects on regional technological specialization, coordinated innovation policies are needed to contrast the detrimental effects of harsh spatial competition processes and to enhance collaboration among the regional innovation systems.

Our study has also shown that the effects could be different according to the way in which the knowledge space is represented, with the top-down macro approach mainly favouring in the long run localized technological change.

The results provided so far could be made more robust by tackling some remaining limits of the current analysis. In future research, our intention is to account for the pronounced asymmetry displayed by the RTA variable and to perform a subsample analysis to deepen our understanding on the evolution of technological specialization according to different characteristics of the regions and of technological macro-classes.

References

Archibugi, D. and Pianta, M. (1992), The Technological Specialization of Advanced Countries. Kluwer, Academic Publishers, Dordrecht.

Acs, Z.J., Anselin, L., Varga, A. (2002), Patents and innovation counts as measures of regional production of new knowledge, *Research Policy*, 31, 1069-1085.

Arthur, W. B. (1990), 'Silicon Valley'locational clusters: when do increasing returns imply monopoly?, *Mathematical Social Sciences*, 19(3), 235-251.

Arthur, B. (1989), Competing technologies, increasing returns, and lock-in by historical events, *Economic Journal*, 116-131.

Atkinson, A B. and Stiglitz, J.E. (1969), A New View of Technological Change, *Economic Journal*, LXXIX, 573-78.

Audretsch, D.B. and Feldman, M.P. (1996), R&D spillovers and the geography of innovation and production, *American Economic Review*, 86, 630-640.

Audretsch, D. and Feldman, M. (2004), Knowledge spillovers and the geography of innovation, in: J.V. Henderson and J. F. Thisse (Eds) *Handbook of Urban and Regional Economics* , Amsterdam: North-Holland.

Autant-Bernard, C., Billand, P., Frachisse, D. and Massard, N. (2007), Social distance versus spatial distance in R&D cooperation: Empirical evidence from European collaboration choices in micro and nanotechnologies, *Papers in Regional Science*, 86, 495-519.

Balland, P.A. (2012), Proximity and the evolution of collaboration networks - evidences from R&D projects within the GNSS industry, *Regional Studies*, 46, 6, 741-756.

Bathelt, H. and Boggs, J. S. (2003), Towards a reconceptualization of regional development paths: Is Leipzig's media cluster a continuation of or a rupture with the past? *Economic Geography*, 79, 265-293.

Boschma, R.A. (2005), Proximity and innovation, A critical assessment, *Regional Studies*, 39, 61-74.

Boschma, R. Balland, A. and Kogler, D. (2014), Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010, *Industrial and Corporate Change*, 1-28.

Boschma, R. and K. Frenken (2010) The spatial evolution of innovation networks. A proximity perspective, in: R. Boschma and R. Martin (Eds.) *The Handbook of Evolutionary Economic Geography*, Cheltenham: Edward Elgar, pp. 120-135.

Boschma, R., and Frenken, K. (2011), Technological relatedness and regional branching. In: H. Bathelt, M. P. Feldman and D. F. Kogler (Eds) *Dynamic geographies of knowledge creation and innovation*. London: Routledge,Taylor and Francis.

Boschma, R.A. and Iammarino, S. (2009), Related variety, trade linkages and regional growth in Italy. *Economic Geography*, 85, 289-311.

Boschma, R., Minondo, A. and Navarro, M. (2012), Related variety and regional growth in Spain, *Papers in Regional Science*, 91, 241-156.

Boschma, R. A. and Wenting R. (2007), The spatial evolution of the British automobile industry. Does location matter? *Industrial and Corporate Change*, 16, 213-238.

Breschi, S. (2000), The geography of innovation: a cross-sector analysis, *Regional Studies*, 34, 213-229.

Breschi, S. and Lissoni, F. (2009), Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows, *Journal of Economic Geography*, 9, 439-468.

Cairncross, F.(1997), *The Death of Distance: How the Communications Revolution Will Change Our Lives*, Harvard Business School Press, Cambridge (Mass).

Cantner, U. and Meder, A. (2007), Technological proximity and the choice of cooperation partner. *Journal of Economic Interaction and Coordination*, 2, 45-65.

Cassi, L., Plunket, A. (2015), Research collaboration in co-inventor networks: combining closure, bridging and proximities. *Regional Studies*, 49, 936-954.

Colombelli A., Kraffts J. and Quatraro F. (2014), The emergence of new technology-based sectors in European regions: A proximity-based analysis of nanotechnology, *Research Policy* 43, 1681-1696.

Cowan, R. and Jonard, N. (2004), Network structure and the diffusion of knowledge, *Journal of Economic Dynamics and Control*, 28, 1557-1575.

Elhorst, J. P. (2014), Spatial econometrics. From cross-sectional data to spatial panels. Heidelberg: Springer.

Engelsman, E.C. and Van Raan, A.F.J. (1992), A Patent-Based Cartography of Technology, *Research Policy*, 23, 1-26.

Essletzbichler J. (2015), Relatedness, Industrial Branching and Technological Cohesion in US Metropolitan Areas, *Regional Studies*, 49, 752-766.,

Feldman, M.P. and Kogler, D. F. (2010), Stylized facts in the geography of innovation, in: B. Hall and N. Rosenberg (Eds) *Handbook of the Economics of Innovation*, pp. 381-410. Oxford: Elsevier.

Feldman, M.P., Kogler, D. F. and Rigby D. L. (2015), rKnowledge: the spatial diffusion and adoption of rDNA methods. *Regional Studies*, 49, 798-817.

Freeman, L.C. (1979), Centrality in Social Networks. Conceptual Clarification, *Social Networks*, 1, 215-239.

Friedman, T.L. (2005), *The World Is Flat*, Farrar, Straus and Giroux, New York.

Glaeser, E. L. (2005), Reinventing Boston: 1630-2003, *Journal of Economic Geography*, 5, 119-153.

Greunz, L. (2003), Geographically and technologically mediated knowledge spillovers between European regions, *Annals of Regional Science*, 37, 657-680.

Griliches Z., (1979), Issues in assessing the contribution of research and development to productivity growth, *Bell Journal of Economics*, 10, 92-116.

Hausmann, R. and Klinger B. (2007), The structure of the product space and the evolution of comparative advantage. Working paper no. 146, Center for International Development, Harvard University, Cambridge, Mass.

Hidalgo, C.A. and Hausmann, R. (2008), A network view of economic development, *Developing Alternatives*, 12, 5-10.

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

Hall, B. H., Jaffe, A. B. and Trajtenberg, M. (2001), The NBER Patent Citations Data File: Lessons, Insights and Methodology Tools, *National Bureau of Economic Research* (Cambridge, MA) Working Paper No. 8498.

Jaffe, A.B., Trajtenberg, M. and Henderson, R. (1993) Geographic localization of knowledge spillovers as evidenced by patient citations, *Quarterly Journal of Economics*, 108, 577-598.

Kelejian H.H. and Prucha I.R. (2010), Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances, *Journal of Econometrics*, 157, 53-67.

Kogler D.F., Essletzbichler J. and Rigby D.L. (2017), The Evolution of Specialization in the EU15 Knowledge Space, *Journal of Economic Geography*, 17, 345-373.

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

Krackhardt D. (1988), Predicting with networks: nonparametric multiple regression analysis of dyadic data, *Social Networks*, 10, 359-381.

Lawson, C. (1999), Towards a competence theory of the region, *Cambridge Journal of Economics*, 23, 151-166.

Leoncini R., Maggioni M.A. and Montresor S. (1996), Intersectoral Innovation Flows and National Technological Systems: Network Analysis for comparing Italy and Germany, *Research Policy*, 25, 415-30.

LeSage J. P. and Pace R. K. (2009). *Introduction to spatial econometrics*. Boca Raton: CRC Press.

Maggioni, M.A. (2002), *Clustering Dynamics and the Location of High-Tech Firms*, Springer Verlag - Heidelberg and New York.

Maggioni M.A., Uberti T.E., and Nosvelli M. (2017), The 'Political' Geography of Research Networks.FP6 within a 'Two Speed' ERA, *International Regional Science Review*, 40, 337-376.

Maggioni M.A., Nosvelli M. and Uberti T.E. (2007), Space vs. networks in the geography of innovation: a European analysis, *Papers in Regional Science*, 86, 471-493.

Maggioni, M.A., Uberti, T.E. and Usai, S. (2011), Treating patents as relational data: knowledge transfers and spillovers across Italian provinces, *Industry & Innovation*, 18, 39-67.

Maggioni M.A., Nosvelli M. and Uberti T.E. (2014), Does intentional mean hierarchical? Knowledge flows and innovative performance of European regions, *Annals of Regional Science*, 53, 453-85.

Miguélez, E. and Moreno, R. (2013), Research Networks and Inventors' Mobility as Drivers of Innovation: Evidence from Europe, *Regional Studies*, 47, 1668-1685.

Miguélez, E. and Moreno, R. (2017), Networks, Diffusion of Knowledge, and Regional Innovative Performance, *International Regional Science Review*, 40, 331-336.

Moreno, R., Paci, R. and Usai, S. (2005), Geographical and sectoral clusters of innovation in Europe, *Annals of Regional Science*, 39, 715-739.

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

Paci R., Sassu A. and Usai S. (1997), International patenting and national technological specialization *Technovation*, 17 , 25-38.

Paci, R. and Usai, S. (2009), Knowledge flows across the European regions, *Annals of Regional Science*, 43, 669-690.

Pavitt K (1988), Sectoral patterns of technical change: Towards a taxonomy and a theory, *Research Policy*, 13, 343-373.

Picci, L. (2010), The internationalization of inventive activity: A gravity model using patent data, *Research Policy*, 39, 1070-1081.

Ponds, R., van Oort, F. and Frenken, K. (2007), The geographical and institutional proximity of research collaboration, *Papers in Regional Science*, 86, 423-444.

Rigby, D.L. (2013), The geography of knowledge relatedness and technological diversification in U.S. cities, Papers in Evolutionary Economic Geography (PEEG) 1218, Utrecht University.

Shaw J. (2010), Complexity and the Wealth of Nations, *Harvard Magazine*, March-April, http://harvardmagazine.com/2010/03/complexity-and-wealth-of-nations, last accessed in July 2017.

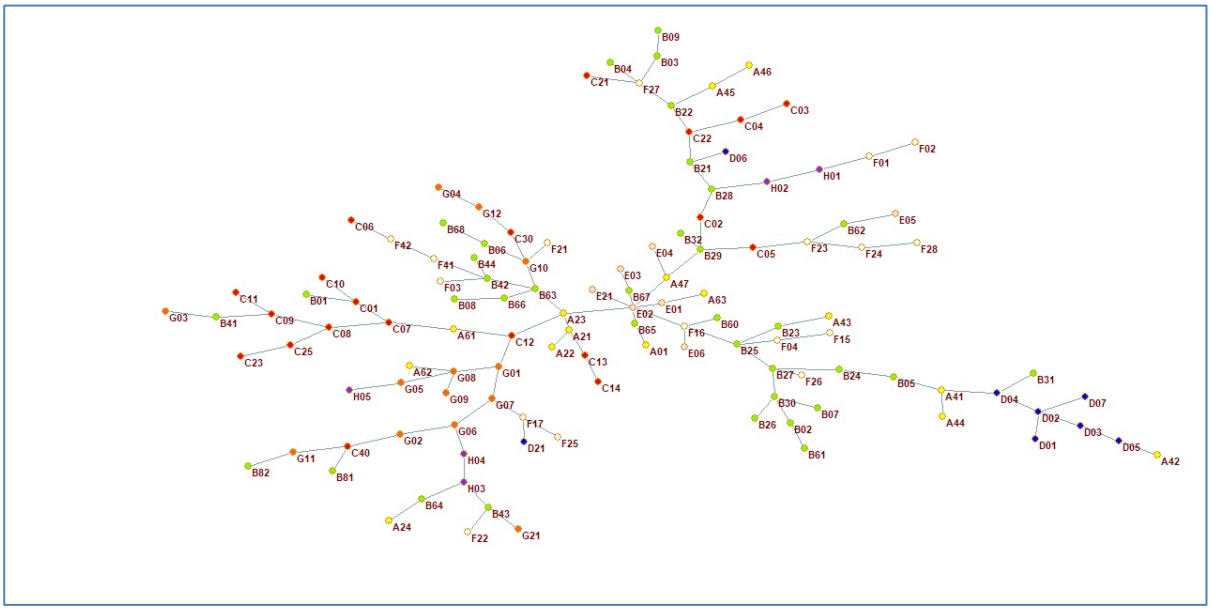
Soete, L. and Wyatt, S. (1983), The use of foreign patenting as an internationally comparable science and technology output indicator. *Scientometrics*, 5, 31.54.

Torre, A. and Gilly, J.P. (2000), On the analytical dimension of proximity dynamics, *Regional Studies*, 34, 169-180.

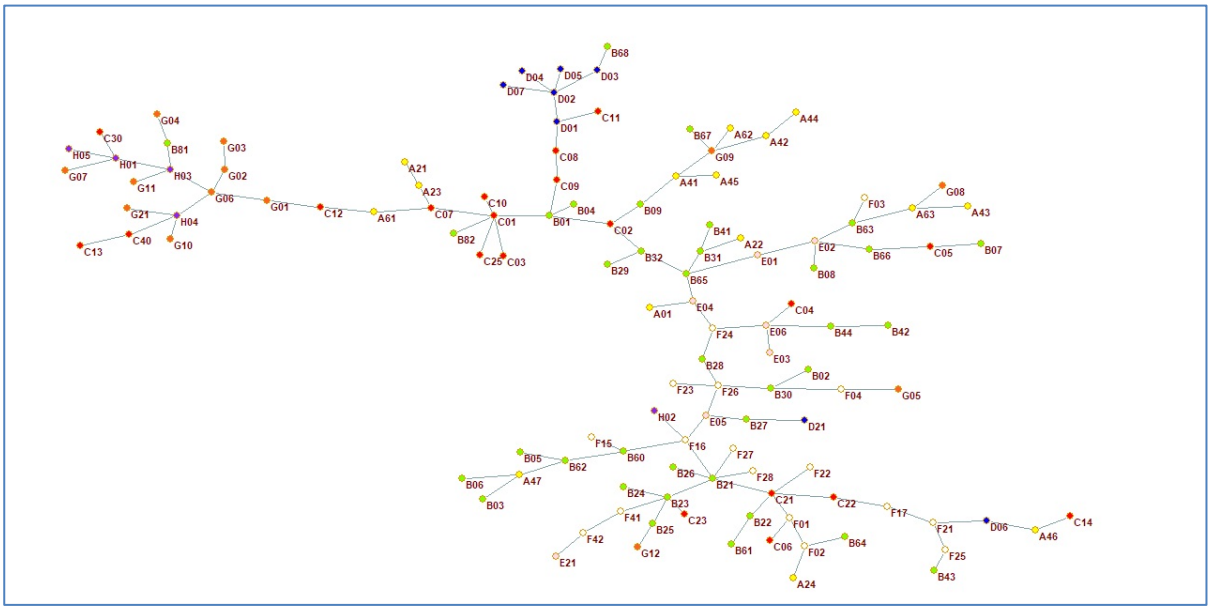
Weitzman M.L. (1998), Recombinant Growth, *The Quarterly Journal of Economics*, 113, 331-360.

FIGURES AND TABLES

Figure 1 – European MST, HH approach



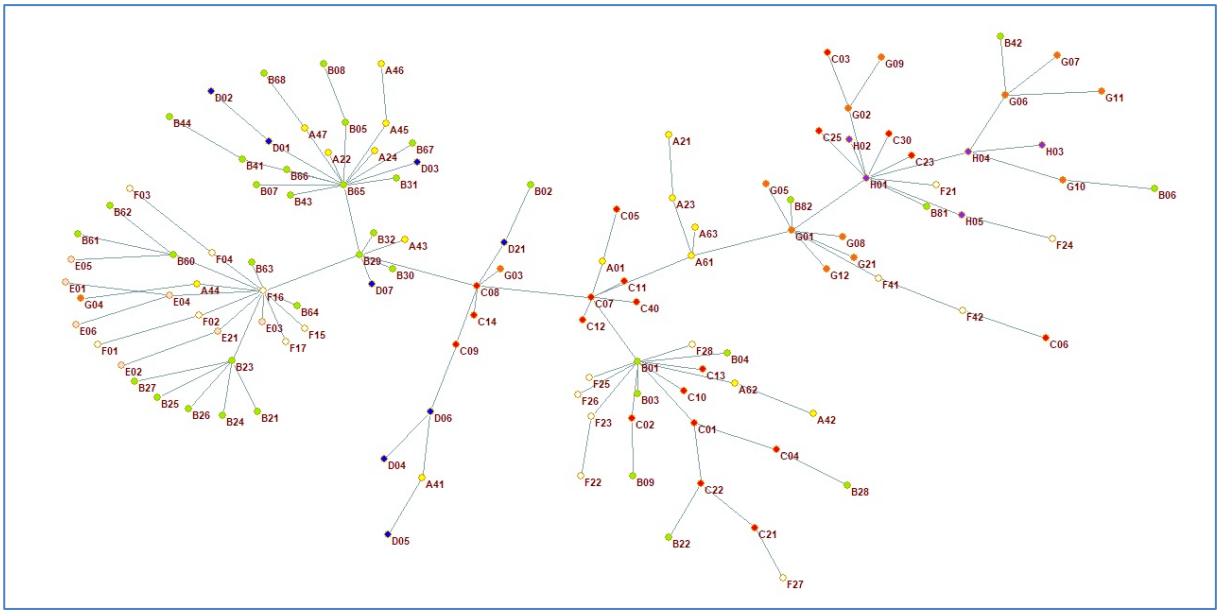
T1: 1986-1990



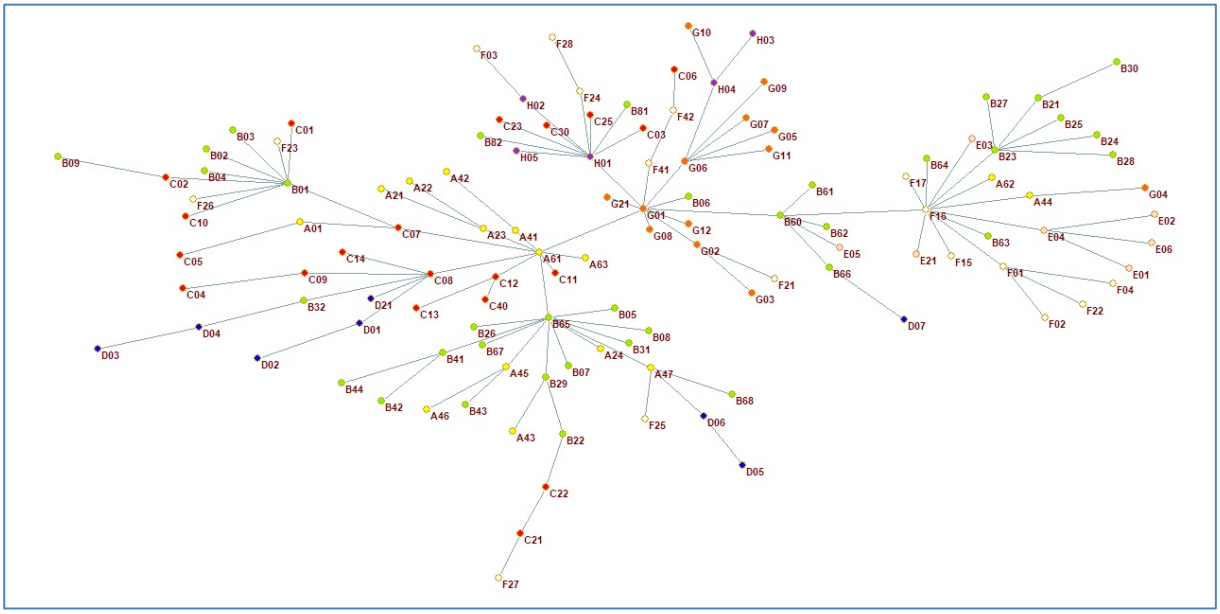
T5: 2005-2010

A: Human Necessities	B: Performing operations; transporting	C: Chemistry; metallurgy	D: Textiles; paper
E: Fixed constructions	F: Mechanical engineering; lighting; heating; weapons; blasting	G: Physics	H: Electricity

Figure 2 – European MST, CC approach



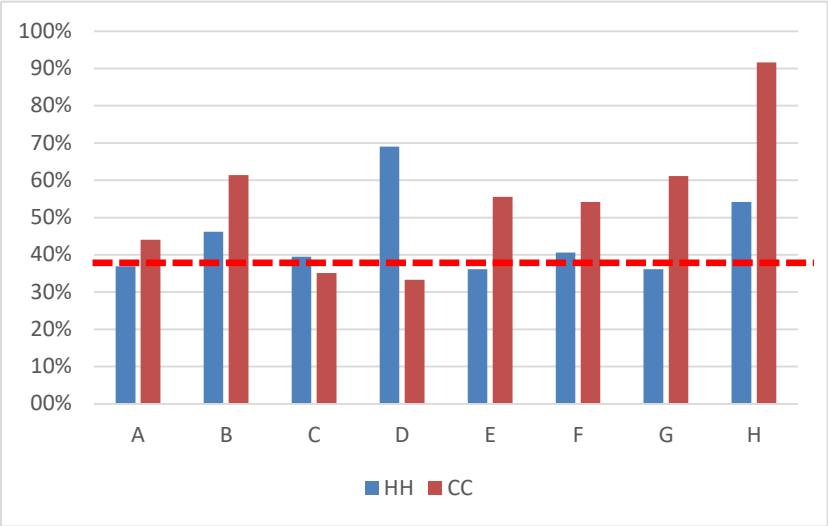
T1: 1986-1990



T5: 2005-2010

A: Human Necessities	B: Performing operations; transporting	C: Chemistry; metallurgy	D: Textiles; paper
E: Fixed constructions	F: Mechanical engineering; lighting; heating; weapons; blasting	G: Physics	H: Electricity

Figure 3 – Average density values for one-digit IPC sectors



Legend: A: Human Necessities (15 IPC classes at second hierarchy level); B: Performing operations; transporting (36); C: Chemistry; metallurgy (20); D: Textiles; paper (8); E: Fixed constructions (7); F: Mechanical engineering; lighting; heating; weapons; blasting (17); G: Physics (13); H: Electricity (5).

Table 1 - QAP correlations for MST - *HH* and *CC* approaches

	HH_t1	HH_t2	HH_t3	HH_t4	HH_t5	HH_t6	CC_t1	CC_t2	CC_t3	CC_t4	CC_t5	CC_t6
HH_t1	*1.000*						***0.195***	0.195	0.204	0.187	0.178	0.170
HH_t2	0.204	*1.000*					0.153	***0.170***	0.161	0.161	0.153	0.153
HH_t3	0.178	0.220	*1.000*				0.136	0.144	***0.153***	0.161	0.153	0.153
HH_t4	0.212	0.220	0.246	*1.000*			0.195	0.220	0.220	***0.220***	0.195	0.195
HH_t5	0.178	0.170	0.187	0.246	*1.000*		0.144	0.136	0.136	0.153	***0.178***	0.178
HH_t6	0.153	0.170	0.170	0.212	0.212	*1.000*	0.170	0.170	0.187	0.204	0.195	***0.195***
CC_t1							*1.000*					
CC_t2							0.754	*1.000*				
CC_t3							0.729	0.797	*1.000*			
CC_t4							0.602	0.644	0.695	*1.000*		
CC_t5							0.619	0.661	0.661	0.763	*1.000*	
CC_t6							0.602	0.619	0.636	0.712	0.831	1.000

Note: All coefficients are statistically significant at 1%.

Table 2 - Dynamic SLX models for RTA

Dependent Variable RTA_{irt} at region-IPC class level

	Model 1		Model 2		Model 3	
	HH	**CC**	**HH**	**CC**	**HH**	**CC**
	coeff.					
RTA_{irt-1}	0.3590 ***	0.3611 ***	0.3564 ***	0.3599 ***	0.3560 ***	0.3593 ***
	(0.0608)	*(0.0611)*	*(0.0612)*	*(0.0612)*	*(0.0612)*	*(0.0612)*
Localized technological change (LTC)	0.0249 ***	0.0203 ***	0.0240 ***	0.0185 ***	0.0238 ***	0.0214 ***
	(0.0044)	*(0.0038)*	*(0.0044)*	*(0.0036)*	*(0.0046)*	*(0.0041)*
NEighbouring Albedo (NEAL)	0.0121 ***	0.0038	0.0123 ***	0.0041 0.11	0.0121 ***	0.0034
	(0.0031)	*(0.0026)*	*(0.0031)*	*(0.0026)*	*(0.0032)*	*(0.0026)*
Recombinant Innovation (RI)	-0.0001	-0.0006 ***	-0.0010 ***	-0.0012 ***	-0.0005 ***	-0.0008 ***
	(0.0001)	*(0.0001)*	*(0.0002)*	*(0.0002)*	*(0.0002)*	*(0.0001)*
RI * DRTA(=1 if RTA>1)			0.0016 ***	0.0016 ***	0.0014 ***	0.0015 ***
			(0.0004)	*(0.0002)*	*(0.0004)*	*(0.0002)*
Exogenous technological shifts - LTC	0.0039 *	0.0053 *	0.0036 0.11	0.0047 0.12	0.0037 0.11	0.0062 *
	(0.0024)	*(0.0031)*	*(0.0023)*	*(0.0030)*	*(0.0023)*	*(0.0033)*
Exogenous technological shifts - NEAL	-0.0011	-0.0008	-0.0011	-0.0007	-0.0010	-0.0010
	(0.0016)	*(0.0013)*	*(0.0016)*	*(0.0013)*	*(0.0016)*	*(0.0012)*
Geographic proximity						
W_G * LTC					-0.0397 ***	-0.0462 ***
					(0.0132)	*(0.0082)*
W_G * NEAL					0.0028	0.0007
					(0.0039)	*(0.0023)*
W_G * RI					-0.0024 ***	-0.0012 ***
					(0.0007)	*(0.0004)*
W_G * (RI*DRTA)					0.0043 ***	0.0028 ***
					(0.0011)	*(0.0008)*
Technological proximity						
W_T * LTC					0.0985 ***	0.0259 ***
					(0.0249)	*(0.0085)*
W_T * NEAL					-0.0155 ***	-0.0073 ***
					(0.0033)	*(0.0014)*
W_T * RI					-0.0006 **	0.0000
					(0.0003)	*(0.0002)*
W_T * (RI*DRTA)					0.0001	0.0005
					(0.0004)	*(0.0004)*
R-squared	0.1625	0.1611	0.1633	0.1617	0.1636	0.1620
Number of observations	95832	95832	95832	95832	95832	95832

Methodology followed to contruct the European knowledge space: HH (Hidalgo et al. (2007), CC (IPC classes co-occurence in patents, Kogler et al., 2015)
Time period: 1985-2010; observations refer to five-year averages
All models include period fixed effects and country dummies
All explanatory variables are lagged one period
W_G: regional geographic proximity matrix (inverse of distance in km); max-eigenvalue normalized
W_T: regional technological connectivity matrix (co-inventorships); max-eigenvalue normalized
Standard errors are reported in parenthesis and clustered at the region level
Level of significance: *** 1%, ** 5%, * 10%

Table 3 - Long run direct and indirect effects

	HH		CC	
	direct effects			
Localized technological change (LTC)	0.0369	***	0.0335	***
	(0.0083)		*(0.0063)*	
NEighbouring Albedo (NEAL)	0.0188	***	0.0053	
	(0.0047)		*(0.0042)*	
Recombinant Innovation (RI)	-0.0008	***	-0.0013	***
	(0.0002)		*(0.0002)*	
RI * DRTA(=1 if RTA>1)	0.0021	***	0.0024	***
	(0.0004)		*(0.0002)*	
Exogenous technological shifts - LTC	0.0058	0.11	0.0097	*
	(0.0037)		*(0.0050)*	
Exogenous technological shifts - NEAL	-0.0015		-0.0016	
	(0.0025)		*(0.0019)*	
	indirect effects due to geographic proximity			
W_G * LTC	-0.0617	***	-0.0721	***
	(0.0222)		*(0.0124)*	
W_G * NEAL	0.0044		0.0012	
	(0.0059)		*(0.0035)*	
W_G * RI	-0.0038	***	-0.0019	***
	(0.0010)		*(0.0006)*	
W_G * (RI*DRTA)	0.0066	***	0.0044	***
	(0.0017)		*(0.0011)*	
	indirect effects due to technological proximity			
W_T * LTC	0.1530	***	0.0404	***
	(0.0350)		*(0.0130)*	
W_T * NEAL	-0.0240	***	-0.0113	***
	(0.0046)		*(0.0020)*	
W_T * RI	-0.0009	*	0.0000	
	(0.0004)		*(0.0003)*	
W_T * (RI*DRTA)	0.0002		0.0007	
	(0.0006)		*(0.0007)*	

The effects are computed from Model 3 in Table 2.

W_G: regional geographic proximity matrix (inverse of distance in km); max-eigenvalue normalized

W_T: regional technological connectivity matrix (co-inventorships); max-eigenvalue normalized

Standard errors are reported in parenthesis and clustered at the region level

Level of significance: *** 1%, ** 5%, * 10%

APPENDIX

Table A1 - Countries and regions		
Country	**Regions**	**EU**
Austria	9	
Belgium	11	
Switzerland	7	no
Germany	38	
Denmark	5	
Finland	5	
France	22	
Italy	21	
The Netherlands	12	
Norway	7	no
Spain	16	
Sweden	8	
United Kingdom	37	
Total	**198**	

Table A2 - Variable definitions and descriptive statistics

Variable	Definition	MST = HH				MST = CC			
		mean	st. dev.	min	max	mean	st. dev.	min	max
Revealed Technological Advantage (RTA)	Proportion of a region's patents in a given IPC class divided by the proportion of European patents in the same IPC class.*	1.083	3.013	0	261.5294				
DRTA	Dummy variable taking value of 1 if RTA>1	0.292	0.455	0	1				
Localized technological change (LTC)	Sum of the RTA values of contiguous sectors in the European Maximum Spannint Tree (HH or CC)	2.135	4.166	0	261.529	2.064	4.214	0	261.529
NEighbouring ALbedo (NEAL)	Sum of the RTA values of second and higher order neighbouring sectors weighted by the geodesic distance in the MST (HH or CC)	17.599	6.686	0	137.697	26.899	8.820	0	159.992
Recombinant Innovation (RI)	Betweenness centrality index in the European MST (HH or CC). The index is computed only for IPCs with RTA>1	40.943	97.593	0	1430.000	24.073	85.490	0	1729.000
RI*DRTA	Interaction term between RI and DRTA	19.826	76.147	0	1430.000	9.829	55.271	0	1729.000
Exogenous technological shift - LTC	Sum of the RTA values at time *t-1* of contiguous sectors in the European Maximum Spanning Tree at time *t* (HH or CC)	2.180	5.688	0	760.428	2.076	5.156	0	760.428
Exogenous technological shift - NEAL	Sum of the RTA values at time *t-1* of second and higher order neighbouring sectors weighted by the geodesic distance in the MST at time *t* (HH or CC)	17.833	7.810	0	380.214	27.337	11.156	0	380.214
W_G * LTC	Geographical lags of the explanatory variables based on geographic proximity. The matrix W_G is the inverse distance matrix, maximum eigenvalue normalized	1.280	1.001	0.077	26.174	1.251	1.634	0.106	37.573
W_G * NEAL		10.502	4.337	1.638	52.149	16.099	6.216	2.618	64.007
W_G * RI		25.842	54.304	0	918.000	15.097	48.535	0	882.000
W_G * (RI*DRTA)		12.541	28.186	0	579.000	6.382	20.799	0	705.000
W_T * LTC	Socio-cognitive lags of the explanatory variables based on technological proximity. The matrix W_T is the regional co-inventorship matrix, maximum eigenvalue normalized	0.414	0.788	0	16.115	0.417	1.061	0	23.310
W_T * NEAL		3.336	4.994	0	48.291	5.212	7.687	0	67.402
W_T * RI		10.364	40.615	0	1190.000	5.733	33.509	0	1230.000
W_T * (RI*DRTA)		4.857	21.545	0	895.000	2.716	17.952	0	842.000

The primary source of the data is the European Patent Office (EPO).

* Europe refers to the countries included in the analysis: Austria, Belgium, Denmark, Germany, Spain, Finland, France, Italy, Netherlands, Sweden, United Kingdom, Switzerland and Norway

Number of observations: 119790=121 IPC classes * 198 NUTS2 regions * 5 periods (time periods are 5 year-averages over the 1986-2010 years)

Brief description on how to construct a Maximum Spanning Tree (MST)

For the sake of simplicity, suppose you have an *N*x*N* matrix, with *N*=10, which represent patent classes in the technological MST.

Within the HH approach, the off-diagonal cells contain the conditional probability of a region having a Revealed Technological Advantage (RTA) > 1 in IPC class *i* given that it has RTA>1 in IPC class *j*. Within the CC approach the off-diagonal elements are the of patents quoting both class *i* and class *j* in the EPO application.

In our study the actual matrix is: 121x121 IPC classes

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	24	24	21	**23**	20	22	**23**	22	21
2	24	0	28	16	18	19	22	19	21	20
3	24	28	0	18	19	18	22	19	25	23
4	21	16	18	0	20	18	19	21	15	15
5	23	18	19	20	0	21	19	23	13	13
6	20	19	18	18	21	0	23	21	14	14
7	22	22	22	19	19	23	0	22	17	17
8	23	19	19	21	**23**	21	22	0	15	14
9	22	21	25	15	13	14	17	15	0	40
10	21	20	23	15	13	14	17	14	40	0

1) First look for the highest non diagonal value in the proximity matrix and draw a link between the identified two nodes

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0	24	24	21	**23**	20	22	**23**	22	21
2	24	0	28	16	18	19	22	19	21	20
3	24	28	0	18	19	18	22	19	25	23
4	21	16	18	0	20	18	19	21	15	15
5	23	18	19	20	0	21	19	23	13	13
6	20	19	18	18	21	0	23	21	14	14
7	22	22	22	19	19	23	0	22	17	17
8	23	19	19	21	**23**	21	22	0	15	14
9	22	21	25	15	13	14	17	15	0	40
10	21	20	23	15	13	14	17	14	40	0

9 — **10**

2) Then, look for the closest node to the identified couple of nodes

	1	2	3	4	5	6	7	8	9	10
1	0	24	24	21	23	20	22	23	22	21
2	24	0	28	16	18	19	22	19	21	20
3	24	28	0	18	19	18	22	19	25	23
4	21	16	18	0	20	18	19	21	15	15
5	23	18	19	20	0					
6	20	19	18	18	21					
7	22	22	22	19	19					
8	23	19	19	21	23					
9	22	21	25	15	13					
10	21	20	23	15	13					

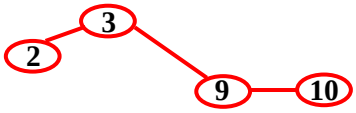
	1	2	3	4	5	6	7	8	9	10
1	0	24	24	21	23	20	22	23	22	21
2	24	0	28	16	18	19	22	19	21	20
3	24	28	0	18	19	18	22	19	25	23
4	21	16	18	0	20	18	19	21	15	15
5	23	18	19	20	0	21	19	23	13	13
6	20	19	18	18	21	0	23	21	14	14
7	22	22	22	19	19	23	0	22	17	17
8	23	19	19	21	23	21	22	0	15	14
9	22	21	25	15	13	14	17	15	0	40
10	21	20	23	15	13	14	17	14	40	0

3) Then, look for the closest node to the identified triplet of nodes

	1	2	3	4	5	6	7	8	9	10
1	0	24	24	21	23	20	22	23	22	21
2	24	0	28	16	18	19	22	19	21	20
3	24	28	0	18	19	18	22	19	25	23
4	21	16	18	0	20	18	19	21	15	15
5	23	18	19	20						
6	20	19	18	18						
7	22	22	22	19						
8	23	19	19	21						
9	22	21	25	15						
10	21	20	23	15						

	1	2	3	4	5	6	7	8	9	10
1	0	24	24	21	23	20	22	23	22	21
2	24	0	28	16	18	19	22	19	21	20
3	24	28	0	18	19	18	22	19	25	23
4	21	16	18	0	20	18	19	21	15	15
5	23	18	19	20	0					
6	20	19	18	18	21					
7	22	22	22	19	19					
8	23	19	19	21	23					
9	22	21	25	15	13					
10	21	20	23	15	13					

	1	2	3	4	5	6	7	8	9	10
1	0	24	24	21	23	20	22	23	22	21
2	24	0	28	16	18	19	22	19	21	20
3	24	28	0	18	19	18	22	19	25	23
4	21	16	18	0	20	18	19	21	15	15
5	23	18	19	20	0	21	19	23	13	13
6	20	19	18	18	21	0	23	21	14	14
7	22	22	22	19	19	23	0	22	17	17
8	23	19	19	21	23	21	22	0	15	14
9	22	21	25	15	13	14	17	15	0	40
10	21	20	23	15	13	14	17	14	40	0



Following the same procedure, given a number of IPC class equal to *N*, we end up with a minimum connected network with *N-1* links.