Paper to be presented at

DRUID15, Rome, June 15-17, 2015

(Coorganized with LUISS)

Core and Periphery in Scientific Networks: Evidence from European Inter-Regional Collaborations, 2000-2011

Lorenzo Cassi
University of Paris 1
Centre Economie de la Sorbonne
lorenzo.Cassi@univ-paris1.fr

Emilie-Pauline Gallié
HCERES
OST
emilie.pauline.gallie@obs-ost.fr

agenor Lahatte
HCERES
OST
agenor.lahatte@obs-ost.fr

Valérie Merindol
ESG Management School
Strategy Management
valerie.merindol@obs-ost.fr

Abstract

This paper provides an original framework for investigating scientific collaboration networks at European regional level. Are the European regions organized around a core? If yes, which are the determinants explaining the core membership? To answer these questions, the evolution of eight different scientific discipline networks from 2000 to 2011 is taken into account. For each of these networks, we identify a core-periphery structure and we investigate the determinants of being in the core in terms of scientific production both in terms of quantity and quality, resource endowments and absorptive capacity. Overall, the scientific production and its quality are the main drivers. The importance of being an affluent region (GDP per capita) and with already a strong reputation in other disciplines (i.e. core in other disciplines) play a major role for determining who can enter the club of the core regions. These two factors work as a reinforcing mechanism and makes be very unlikely the possibility to overcome the core-periphery structure. Finally, results obtained confirm the interest of analysing the centrality discipline by discipline rather than globally and to take into account spatial effects.

Jelcodes:R12,-

Core and Periphery in Scientific Networks:
Evidence from European Inter-Regional Collaborations, 2000-2011

February 2015

This paper provides an original framework for investigating scientific collaboration networks at European regional level. Are the European regions organized around a core? If yes, which are the determinants explaining the core membership? To answer these questions, the evolution of eight different scientific discipline networks from 2000 to 2011 is taken into account. For each of these networks, we identify a core-periphery structure and we investigate the determinants of being in the core in terms of scientific production both in terms of quantity and quality, resource endowments and absorptive capacity. Overall, the scientific production and its quality are the main drivers. The importance of being an affluent region (GDP per capita) and with already a strong reputation in other disciplines (i.e. core in other disciplines) play a major role for determining who can enter the club of the core regions. These two factors work as a reinforcing mechanism and makes be very unlikely the possibility to overcome the core-periphery structure. Finally, results obtained confirm the interest of analysing the centrality discipline by discipline rather than globally and to take into account spatial effects.

Introduction

Since the Lisbon strategy, regions are at the heart of European policy in order to build the Europe of knowledge. The variety of regional policies and of European instruments (e.g. framework programmes and cohesion policy) is a consequence of the diversity of regional situations in terms of their capabilities in R&D and innovation (Asheim, 2006; Pinto, 2009). This diversity is frequently presented as a dichotomy between core/leader vs. periphery/follower (Winthjes et al., 2010; OECD, 2011). The key difference between these two types of regions is based on two criteria. The first is the ability to concentrate a wide variety of industrial, technological and scientific resources allowing regions to achieve critical mass and to develop enough absorptive capacity in order to participate in the dynamics of innovation in Europe (Harris et al., 2005; Almeida et al. 2011). The second criterion is based on the differences in innovation performance (European Commission, 2012). The difference in performance is based on both the lack of resources (critical mass) and the inability of regional systems to transform existing knowledge on the territory value (Cooke, 2001).

In this perspective, the peripheral regions are a specific target of public policies. With low capabilities and low performance, peripheral regions are often not well integrated into the European research area. For policymakers, the challenge is therefore to develop devices that allow peripheral regions to move out of their isolation (Doloreux et al., 2008; Hewitt- Dundas et al., 2011). If the study of scientific capacity in peripheral regions has been an object of analysis barely explored, it shares questions similar to those addressed by regional systems literature. The same dichotomy based on the same criteria of quantity and quality: regional differences can be analysed in terms of critical mass, i.e. scientific production, and performance in terms of scientific excellence, i.e. impact or visibility (Zitt et al., 2000; Tjissen et al., 2007). The consequence of such approaches to regional disparities is that policy actions are designed to increase the scientific capacity of research institutions on the one hand and to push to pursuit of scientific excellence on the other hand (Harris et al., 2005; Benneworth et al. 2007).

With the new EU initiatives developed in 2020, the strengthening of regional scientific capabilities is no longer based only on the resource/activities concentration (e.g. critical mass - absorption capacity) and performance targets (e.g. excellence, reputation). It is also based on another dimension that should be taken into account together with the other two: dynamic networks (Winthjes et al., 2010). In fact, in 2020, the pursuit of regional *pôle* of scientific excellence is based both on traditional criteria of quantity and scientific reputation (i.e. quality) but also on the ability to build dynamic networks and collaborations within the European research area (European Commission, 2010). The Horizon 2020 strategy modifies priorities considering the dynamic networks central for developing centres of excellence. The relation between the couple critical mass-excellence and network becomes the main feature of every EU policy. The aims are therefore to develop centres of excellence based on their ability to connect a wide variety of actors and to increase connection between core and peripheral regions.

If the insertion of peripheral regions in research networks is central, it remains unsolved to know how regions that are on the fringes of scientific activities can develop partnerships and move towards the core of European research area. By addressing this question, the aim of this article is to better analyse the dynamic relationships between integration in networks, critical mass and absorptive capacity of regions and their scientific reputation. To do so, we analyse the scientific network in eight different disciplines over four periods of 3 years each, checking whether they are actually structured around core. Moreover, we investigate the determinants of being in the core in terms of quantity and quality of the research carried on by each region, controlling for spatial effects of neighbouring regions.

The paper is organized as follows: the second section presents briefly the empirical literature dealing with similar issues, highlighting the originality of our approach. The third one presents the data and discusses the main assumptions relate to network construction and explains the methodology adopted to detect a core-periphery structure. The following one presents the empirical exercise and the main results. The last section concludes the discussion.

Collaboration network in Europe

The literature dealing with research collaboration among European regions is the main reference of our paper. This literature can be easily classified, among other criteria, according to data used as a proxy of collaboration *and* according to unity of analysis. Usually three kinds of data are analysed: research project (e.g. Framework Program), co-patent and scientific co-publication (for a review see Frenken et al., 2009). Concerning the unity of analysis, the literature focuses, on one hand, on individual actors (i.e. regions) analysing the position of each of them in collaboration networks or their collaboration pattern and, on the other hand, on the overall structure of network. The former is the local scale perspective, the latter the global one. Less studied is the intermediate-scale of EU research collaboration network. In this paper, using publication data, we investigate the meso-scale network feature known as core-periphery structure, which consists of a partition of network's nodes in a highly connected core and a sparsely connected periphery (Rombach et al., 2012).

Several scientific studies have examined the impact of the variety of collaborations and the position of actors in networks. The actors having a central position in a network benefit of great advantages (Wanzenbock et al., 2014). Central positions allow regions developing their own capacity in terms of research institutions because of their greater influences and easier access to knowledge and information, and financial support. For peripheral regions, networks constitute a challenge but also a constraint: peripheral regions are both areas that are struggling to position themselves more centrally, and at the same time they necessarily need to rely on dynamic external interactions to promote their development and their ability to achieve the critical mass that is necessary (Doloreux et al. 2008). Integration into networks is a particularly critical challenge for development and at the same time the initial endowments and capacities developed are required to strengthen their presence in a dynamic network.

What we intend to investigate are the determinants of being member of the core, that means a highly central region mainly connected with other region having a similar centrality.

The factors we are going to consider concerns the scientific production of each region, both in terms of critical mass (quantity) and excellence (quality). Moreover, we are going to take into account the region resource endowments and its absorptive capacity. The spatial effect will be considered as well in terms of spillovers generated by neighbouring regions.

Regional networks: data and methods

Data

In order to investigate the European scientific system we focus on the collaboration existing among regions (NUTS2) in eight broad scientific disciplines (see Table 1) that are defined by OST (2010) as an aggregation of Thomson Reuters Scientific Categories. Moreover, those disciplines are completed by reallocating the articles published in multidisciplinary journals.

Discipline
Fundamental biology
Medicine
Applied biology/ecology
Chemistry
Physics
Science of the universe
Engineering sciences
Mathematics

Table 1: The scientific disciplines.

Data of co-publications among NUTS2 regions (EU27) used and citations associated come from Web of Science (WoS) database, which contains information from most journals covering all scientific fields. We retrieve all scientific articles, letters, notes and reviews published between 2000 and 2011 and related to research collaborations among EU regions. A publication is considered to be research collaboration between regions if it contains at least two different institutional addresses corresponding to NUTS2 regions. Consequently the publications co-authored by intra-regional institutions are excluded and our study is limited to the bilateral and multilateral inter-regional co-authorship and is in fact performed on four sets smoothed data corresponding to 3-years periods: 2000-2002, 2003-2005, 2006-2008, and 2009-2011. Our empirical study covers 265 regions across 27 European countries.

Methods

We use the data on scientific publications co-authored by inter-regional institutions to build a network, by discipline and by period, where the actors/nodes are the regions and the links between them are co-publications. The value of the link between two regions is given by the number of collaborations in common in a given discipline over the period under analysis. That means we are considering a weighted network.

Following the methodology initially developed by Zhou and Mondargon (2003) and generalised by Serrano (2008), we identify a core and a periphery according to the rich-club phenomenon. In order to do that, we analyse the distribution of degree among regions for each network associated to each discipline and period. However, given the fact we are considering a weighted network, referring to the degree centrality can be a tricky issue as emphasised by Opsahl et al. (2010).

Indeed, in a binary network (i.e. not weighted network), the centrality degree of an actor is the number of link of the actor: in our case it would be the number of regions which the focus regions is collaborating with. In a weighted network, the degree centrality is given by the sum of the weight of actor's links: in our case it would correspond to the number of inter-regions collaborations carried on by the focal region. However in the latter case, we lost completely the information about the number of partners: the same value of centrality (e.g. ten) can correspond to very different scenarios (e.g. one partner with which you have collaborated ten times or ten collaborations with ten different partners). Opsahl et al. 2010 propose a definition of the degree in weighted network that is able to grasp both aspects of collaboration: quantity (i.e. degree) and strength. The weighted centrality degree for actor *i* is therefore defined as the following:

$$C_D^{w\alpha}(i) = k_i \times \left(\frac{s_i}{k_i}\right)^{\alpha} = k_i^{(1-\alpha)} \times s_i^{\alpha} \tag{1}$$

where k_i is the number of partners and s_i the sum of actor's links strength, and α is the parameter weighting each of the two aspects. Following XXXX (20XX), we decide to set α equal to 0.5 because we have no specific reasons to assign more importance to one aspect rather than the other.

Once calculated the degree centrality of each region, $C_D^{w\alpha}(i)$, we look at the characteristic of its distribution in order to identify the existence of core-periphery structure (Serrano, 2008). In order to establish whether a network is characterized by such a similar structure, a necessary condition is the presence of break-point in the distribution. To check for that, we (i) compute the log-CCDF (complementary cumulative distribution function) of log-degree; (ii) "use a piece-wise linear regression of log-degree and log-CCDF using a moving threshold to get best fit for power law coefficients" (Zelnio 2012, p. 604).

If this point is identified, then it is possible to distinguish the regions in central and peripheral ones. However, the most central regions constitute a core, only if the links between them are particularly intensive, i.e. the number of links between core members is greater than "the corresponding value in a randomized version of the graph that preserves the degree distribution" (Serrano 2008, p.4). Once this second condition is satisfied, it is possible to claim that the network analysed displays a core-periphery structure.

Network and core identification

We build 32 networks: eight discipline/networks for each of the four periods. Table 2 reports the main information about the distribution links value, i.e. s_i of equation 1.

Discipline		# Pairs of regions	Mean	Median	Max
	2000-2002				
Fundamental biology		12106	3.041	1.000	236.333
Medicine		13728	5.031	1.000	390.333
Applied biology/eco		7670	1.463	0.667	61.667
Chemistry		9897	2.167	0.667	183.000
Physics		10006	6.191	1.333	358.333
Science of the univ		8858	2.150	0.667	118.667
Engineering science		8739	1.726	0.667	127.667
Mathematics		3820	0.839	0.333	22.000
	2003-2005				
Fundamental biology		13270	3.342	1.000	243.000
Medicine		16197	6.072	1.333	534.667
Applied biology/eco		9233	1.579	0.667	77.000
Chemistry		10757	2.307	1.000	211.667
Physics		11063	7.202	1.333	429.667
Science of the univ		10775	3.272	1.000	198.333
Engineering science		10336	2.041	0.667	165.667
Mathematics		4504	0.941	0.333	38.333
	2006-2008				
Fundamental biology		15286	4.040	1.000	297.333
Medicine		18161	7.851	1.333	699.000
Applied biology/eco		11863	1.881	0.667	94.333
Chemistry		12053	2.527	1.000	265.333
Physics		12638	7.836	1.667	517.667
Science of the univ		13243	4.906	1.000	303.333
Engineering science		11558	2.157	0.667	218.000
Mathematics		5301	1.054	0.667	50.000
	2009-2011				
Fundamental biology		17745	5.341	1.333	393.333
Medicine		20729	10.145	2.000	1009.667
Applied biology/eco		14881	2.418	1.000	136.667
Chemistry		13007	2.789	1.000	322.333
Physics		13507	13.216	2.000	585.333
Science of the univ		16211	6.782	1.667	419.000
Engineering science		12428	2.418	1.000	259.667
Mathematics		6060	1.152	0.667	70.667

Table 2: Scientific networks: weighted link

We have applied the methodology described in previous section to the 32 networks in order to detect a core-periphery structure. All of them result to satisfy the two conditions (results are available on request to the authors) and thus a core-periphery structure is verified. However, as reported in the following table (Table 3), the size of the core, and thus also that of the periphery, varies across disciplines and over time. This could happen because some regions move from the periphery to the core or the other way round.

Discipline		Min degree	Core size	Core share
	2000-2002			
Fundamental biology		191.053	84.000	0.317
Medicine		279.564	88.000	0.332
Applied biology/eco		79.297	83.000	0.313
Chemistry		141.539	79.000	0.298
Physics		312.217	52.000	0.196
Science of the univ		119.733	64.000	0.242
Engineering science		113.287	74.000	0.279
Mathematics		28.337	77.000	0.291
	2003-2005			
Fundamental biology		214.327	84.000	0.317
Medicine		360.342	88.000	0.332
Applied biology/eco		99.465	83.000	0.313
Chemistry		159.499	79.000	0.298
Physics		387.457	52.000	0.196
Science of the univ		197.221	64.000	0.242
Engineering science		140.670	74.000	0.279
Mathematics		36.469	77.000	0.291
	2006-2008			
Fundamental biology		269.560	84.000	0.317
Medicine		433.188	88.000	0.332
Applied biology/eco		134.005	83.000	0.313
Chemistry		195.141	79.000	0.298
Physics		407.113	52.000	0.196
Science of the univ		274.748	64.000	0.242
Engineering science		166.833	74.000	0.279
Mathematics		46.733	77.000	0.291
	2009-2011			
Fundamental biology		348.718	84.000	0.317
Medicine		555.608	88.000	0.332
Applied biology/eco		181.083	83.000	0.313
Chemistry		206.753	79.000	0.298
Physics		612.975	52.000	0.196
Science of the univ		366.999	64.000	0.242
Engineering science		184.748	74.000	0.279
Mathematics		51.020	77.000	0.291

Table 3: Caracteristics of core

The critical threshold of the core represented by the minimum degree increases over time for each domain. Those both elements reflect the network dynamics in different fields studied. The differences among discipline are striking. For instance, the minimum degree for Physics is ten times that for Mathematics, with the other six discipline reporting value between the two extremes. Yet, these descriptive statistics allow understanding that these differences should be taken into account and raise doubts over of analysis covering all disciplines together.

Table 4 reports the density of relation within the core and periphery and between them. Network density is defined as the number of effective links over the potentials one. In undirected networks as those under analyses, the potential links are given by the following formula: N (N-1) / 2 where N is the number of nodes. This can be reformulated according any nodes partition.

Consequently the overall potential links can be broken-down in core (C) and periphery (P) potential nodes:

$$\frac{N(N-1)}{2} = \frac{C(C-1)}{2} + \frac{P(P-1)}{2} + PC \qquad (2)$$

where C+P=N. The three terms are from left to right the potential links: within the core, within the periphery and between core and periphery. The corresponding density can be easily calculated, dividing the effective link with the relative potential links term.

Discipline		Core				Periphery				Core/Periphery			
		Mean	Median	Max	Density	Mean	Median	Max	Density	Mean	Median	Max	Density
	2000-2002												
Fundamental biology		199.922	186.823	448.637	0.960	15.421	12.503	51.607	0.101	9.176	9.292	14.944	0.505
Medicine		281.428	258.766	652.761	0.976	23.731	20.808	73.892	0.126	11.464	10.954	16.155	0.709
Applied biology/eco		93.596	87.651	213.292	0.766	4.924	3.936	22.657	0.040	8.145	7.409	17.436	0.290
Chemistry		130.279	119.304	310.957	0.884	12.603	10.886	35.506	0.077	13.217	12.770	23.087	0.516
Physics		243.325	223.486	398.259	0.995	27.241	21.564	125.037	0.125	25.937	28.460	38.764	0.691
Science of the univ		130.299	119.350	301.232	0.927	8.104	6.325	28.752	0.062	17.097	17.312	25.113	0.497
Engineering science		115.499	105.986	263.344	0.887	9.522	7.415	35.651	0.067	7.767	6.782	14.967	0.419
Mathematics		38.984	35.308	141.772	0.487	1.901	1.155	8.165	0.017	2.796	2.582	5.657	0.129
	2003-2005												
Fundamental biology		218.884	199.652	471.430	0.975	19.650	17.539	65.115	0.122	12.056	11.972	18.974	0.676
Medicine		335.286	316.927	742.723	0.992	35.957	33.510	94.488	0.178	15.621	17.739	21.756	0.836
Applied biology/eco		107.815	97.871	231.459	0.821	7.456	6.506	34.598	0.059	10.310	10.749	17.654	0.345
Chemistry		142.555	133.903	322.006	0.906	15.049	11.626	41.553	0.090	14.309	14.352	23.770	0.614
Physics		282.994	278.087	461.953	1.000	32.396	24.152	139.666	0.143	29.494	30.216	45.129	0.787
Science of the univ		180.547	169.771	394.980	0.977	12.755	10.132	44.796	0.090	28.861	26.745	58.634	0.606
Engineering science		136.985	123.539	308.729	0.925	13.709	11.314	51.949	0.093	11.026	10.488	25.534	0.537
Mathematics		47.628	39.332	170.738	0.554	2.572	2.000	10.312	0.022	3.624	3.289	7.874	0.159
	2006-2008												
Fundamental biology		259.776	233.198	558.963	0.989	26.467	23.814	83.658	0.156	16.099	16.583	21.726	0.800
Medicine		401.212	369.004	884.693	0.995	49.996	46.303	140.921	0.228	18.575	21.276	23.094	0.855
Applied biology/eco		138.114	131.042	288.164	0.908	11.314	9.658	52.802	0.084	17.179	16.207	27.423	0.536
Chemistry		159.884	149.987	360.952	0.941	18.866	16.415	61.417	0.109	15.602	15.593	24.276	0.634
Physics		315.444	322.095	539.450	0.998	39.365	30.055	154.713	0.176	36.889	39.463	51.039	0.919
Science of the univ		241.242	217.104	513.473	0.992	21.319	16.971	77.610	0.132	43.104	34.078	100.300	0.686
Engineering science		149.808	134.092	341.226	0.946	16.843	14.479	56.125	0.107	12.897	12.138	28.740	0.676
Mathematics		57.760	50.481	188.930	0.623	3.298	2.582	13.663	0.029	4.607	3.826	8.641	0.192
	2009-2011												
Fundamental biology		321.001	294.738	682.756	0.999	38.248	36.260	105.167	0.209	23.300	23.958	33.744	0.962
Medicine		484.651	441.728	1055.520	0.999	71.200	65.616	188.507	0.299	26.147	30.056	31.990	1.000
Applied biology/eco		181.569	173.111	354.319	0.965	17.689	16.672	64.410	0.125	23.077	22.519	32.650	0.655
Chemistry		176.613	167.911	386.619	0.954	22.375	19.946	82.049	0.125	15.378	15.045	24.542	0.627
Physics		400.218	386.418	624.661	1.000	54.985	37.722	203.386	0.197	61.799	60.564	75.011	0.912
Science of the univ		318.481	299.611	677.431	0.999	36.393	28.542	119.001	0.196	57.847	53.706	111.310	0.806
Engineering science		162.102	145.193	366.098	0.956	20.004	17.846	63.807	0.125	15.752	12.806	34.583	0.684
Mathematics		65.815	59.944	221.519	0.668	4.104	3.936	11.314	0.036	5.376	4.865	17.550	0.208

Table 4: core/periphery links

As expected, the density of the core is greater than that of periphery and moreover peripheral regions are on average more connected to core regions than with each other. This is true for all disciplines. However, differences among disciplines are remarkable.

The determinants of being in the core

The aim of our empirical exercise is at identifying driving factors of region position in the inter-regional scientific co-publications network. What does it explain the fact that a region belongs to core rather than to the periphery in a given discipline? In order to answer to this question and to compare results over different disciplines, we refer as dependent variable the degree centrality of the region normalized relative to the minimum degree necessary to belong to the core. This normalization allows us comparing the results in terms of magnitude of the coefficients otherwise it would be meaningless.

Moreover, we intend to control for spatial effect, in particular to check if the regions close to region *i* affect its possibility to belong to the core. To do that, we use an analytical framework accounting for spatial knowledge spillover effects in our co-publications network analysis of EU regions (Autant-Bernard, 2012). A panel version of spatial Durbin model (Elhorst 2003 and 2012), Wanzenbock et al. (2014) is applied to our degree centrality data. This model turns out to be more appropriate in empirical analysis for calculating the magnitude of direct impacts and indirect or spatial spillover effects (Autant-Bernard and LeSage, 2011 and Elhorst, 2012). In our study context it allows to distinguish fields with local and? indirect impacts and those with global spillover effects.

The panel SDM is written as follows

$$y_t = \rho W y_t + X_t \beta_1 + W X_t \beta_2 + u_t \quad (3)$$

$$u_t = \mu + \tau_t + \epsilon_t \quad (4)$$

where y_t is an *Nx1* vector containing degree centrality in a discipline normalized relative to the minimum degree for belonging to the core, for each region (*i=1,..., N*) (Table 3) at time *t* where *t=1, 2, 3, 4*. X_t is an *NxK* matrix of exogenous explanatory variables potentially including a constant term. Over all period under observation, the data on publications of regions in each discipline, their total scientific production are taken into account. The relative impact of each region output in each domain is considered. This is measured by the average number of citations received by the region in a discipline given benchmarked against the average number of citations received by the EU in the same domain. Moreover we consider some economic data for the period 1995-2011 retrieved from the Eurostat website in June 2013. In particular we have downloaded for each region: the population (thousands of people); the human resources in Science and Technology, broadly defined (thousands of people); the GNP per capita relative to EU27 average, normalised to 100; the R&D expenditures intensity (% of GDP), broken-down by Public and Private sector. To overcome a minor problem of missing data for some regions in collected Eurostat data, we use the median ratio procedure for estimating these missing values as suggested by the methodology report of Regional Innovation Scoreboard (Hollanders et al., 2012). The number of disciplines which a region is already in the core is also included in the specification as independent variable for underlining a potential effect of core memberships in others disciplines. *W* is an *NxN* nonnegative matrix of known constants describing the connections between regions. *W*, generically labelled spatial weight matrix (see, e.g., Anselin 1988) is constant over time, with the element of *W* in row *i* and column *j* denoted by *w{ij}*. The components of *W* are given by: *w{ij}=0* for all i=j by assumption since no spatial unit be considered as its own neighbour, and *w{ij}=1* if region i and region j are contiguous and *w{ij}=0* otherwise. *W* is row-standardised, meaning that the row elements sum up to 1. Wy_t

corresponds to *Nx1* vector of the dependent variable lagged in space and WX_t is a *NxK* spatial lag matrix of *K* independent variables. The centrality of each region is then supposed rely on? weighted average centrality of its neighbouring regions and to a weighted average of its neighbours' exogenous explanatory variables. The parameter rho associated with spatial lag of y_t is the spatial autoregressive coefficient and reflects the strength of the spatial interaction; β_1, β_2 denote *Nx1* vectors of response parameters of predictors *X* and their spatial lag respectively. u_t denotes a *Nx1* vector of error terms, with μ region specific fixed effects, τ_t period specific fixed effects and ε_t an *Nx1* vector of normally distributed, homoscedastic and uncorrelated errors. The Maximum Likelihood procedure is applied to the spatial Durbin model (Elhorst, 2003) for estimating regression parameters.

Following Wananzenböck et al. (2014), Table 5a and 5b present average impact estimates for each discipline on the magnitude and significance of direct, indirect and total impacts on a region network position that would arise from a change in one unit of our regional characteristics, averaged over space and time. More specifically, the direct impacts give the effects of region-internal characteristics on a region's network centrality, while the indirect impacts estimates report the sum of spatial spillover effects, i.e. influences of region-external characteristics setting. Then, the overall influence of distinct characteristics on regional network centrality at the regional level is given by the total impact estimates.

First, it would worth remarking that geographical space matters in explaining the core membership in five disciplines out the eight analysed. The estimates for the spatial autoregressive parameter (i.e. *rho*) are highly significant and positive for Fundamental Biology, Medicine, Applied Biology, Chemistry and Engineering Sciences but not for Physics, the Science of Universe and Mathematics.

The first column of Table 5a presents impact measures for Fundamental Biology. Concerning direct impact, we observe that a region's capacity in knowledge production is decisive for a region's centrality. We note that both sectorial knowledge production (production in Fundamental Biology) and general knowledge production (public R&D expenditures and being core in other disciplines) account for. The quality of its scientific production (relative impact) plays also a positive role. Direct impact estimates suggest that the regional production in the field and public R&D expenditure are both significantly important capacity factors for being in the core in Fundamental Biology. However, we observe none spatial spillover impacts on a region's network centrality. This result differs from findings in previous literature on the spatial dimension of knowledge production which consider the existence of spatial spillovers.

The total impact estimation confirms direct impact. Moreover, being in a rich region or surrounding by rich regions, also positively influence the region's centrality. Total impact estimates suggest that activities in the field and the level of public R&D expenditures particularly influence network positioning in this discipline. These results suggest that in Fundamental Biology, in order to belong to the core, regions need a large amount of scientific resources.

Table 5a: ML estimation results

	FundBio		Med		ApBioEco		Chem	
	b	t	b	t	b	t	b	t
Main								
GNP per capita	0.000816	0.834	0.00231*	2.471	0.00342***	3.571	0.000678	0.754
Number Fields Core Membership	0.0243***	4.864	0.0158***	4.250	0.0507***	4.682	0.0324***	4.999
Human resources in S& T	-0.000183	-1.067	0.0000553	0.488	-0.0000588	-0.348	-0.000290**	-2.649
Public R& D expenditures	0.175**	3.189	0.0992*	2.245	0.101	1.400	0.121*	2.232
Private R& D expenditures	0.0118	0.605	0.0530**	2.764	0.0339	1.006	0.0228	0.929
Relative Impact	0.0927***	3.999	0.0421***	3.826	0.0767***	3.978	0.0263	1.788
Publication Share by field	0.588**	3.070	0.304***	4.263	0.904***	4.359	0.477***	3.713
Wx								
GNP per capita	0.00189	1.615	0.00143	1.472	0.00304*	2.145	0.000773	0.662
Number Fields Core Membership	0.0150	1.422	0.00136	0.136	-0.00461	-0.329	0.0262**	2.642
Human resources in S& T	0.000160	1.192	0.0000417	0.369	0.000325*	2.175	0.000108	1.084
Public R& D expenditures	0.0775	0.975	-0.0268	-0.462	-0.0313	-0.317	-0.0188	-0.228
Private R& D expenditures	0.0376	0.948	0.0429	1.119	0.0516	0.783	0.0571	1.492
Relative Impact	0.0493	1.223	0.00141	0.063	-0.0330	-0.807	-0.00149	-0.049
Publication Share by field	0.108	0.374	-0.0491	-0.361	-0.102	-0.302	-0.704***	-3.777
Spatial								
rho	0.103*	2.237	0.348***	6.325	0.279***	7.520	0.195***	4.948
Variance								
sigma2_e	0.00378***	9.573	0.00245***	11.105	0.00855***	10.318	0.00386***	12.499
Direct								
GNP per capita	0.000902	0.941	0.00253**	2.860	0.00372***	4.033	0.000753	0.855
Number Fields Core Membership	0.0244***	4.918	0.0162***	4.269	0.0507***	4.639	0.0335***	5.172
Human resources in S& T	-0.000175	-1.017	0.0000617	0.535	-0.0000363	-0.215	-0.000285**	-2.589
Public R& D expenditures	0.175**	3.211	0.0985*	2.231	0.0989	1.375	0.120*	2.225
Private R& D expenditures	0.0126	0.632	0.0587**	2.936	0.0384	1.099	0.0260	1.026
Relative Impact	0.0933***	4.053	0.0432***	3.752	0.0755***	3.896	0.0261	1.762
Publication Share by field	0.582**	3.104	0.305***	4.097	0.903***	4.359	0.443***	3.482
Indirect								
GNP per capita	0.00207	1.643	0.00316*	2.521	0.00516**	2.879	0.000988	0.741
Number Fields Core Membership	0.0194	1.729	0.0100	0.718	0.0129	0.675	0.0391**	3.281
Human resources in S& T	0.000156	1.088	0.0000844	0.524	0.000408*	2.284	0.0000641	0.546
Public R& D expenditures	0.106	1.257	0.0117	0.141	-0.00142	-0.011	0.00646	0.067
Private R& D expenditures	0.0419	0.995	0.0902	1.604	0.0811	0.941	0.0734	1.654
Relative Impact	0.0656	1.483	0.0236	0.691	-0.0141	-0.259	0.00517	0.140
Publication Share by field	0.170	0.548	0.0781	0.397	0.185	0.410	-0.744***	-3.373
Total								
GNP per capita	0.00297***	3.357	0.00569***	5.214	0.00889***	5.270	0.00174	1.382
Number Fields Core Membership	0.0438***	3.500	0.0262	1.716	0.0636**	2.611	0.0727***	5.002
Human resources in S& T	-0.0000191	-0.099	0.000146	0.668	0.000371	1.591	-0.000221	-1.322
Public R& D expenditures	0.282**	2.990	0.110	1.079	0.0975	0.636	0.126	1.133
Private R& D expenditures	0.0545	1.301	0.149*	2.364	0.119	1.235	0.0994	1.959
Relative Impact	0.159***	3.297	0.0668	1.704	0.0614	1.053	0.0313	0.766
Publication Share by field	0.752*	2.206	0.383	1.655	1.089*	2.118	-0.301	-1.200
Observations	1060		1060		1060		1060	

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

The second column of the Table 5a presents impact measures for Medicine. Every variable except the Human Resources in Science and Technology has a significant positive impact. Direct impact estimates suggest that the regional production in the field is significantly an important capacity factor for being embedded in Medicine.

Table 5b: ML estimation results

	Phys		ScUniv		EngSci		Maths	
	b	t	b	t	b	t	b	t
Main								
GNP per capita	0.00183*	2.248	0.00125	0.854	0.00167	1.774	-0.000410	-0.261
Number Fields Core Membership	0.0455***	4.437	0.0810***	4.821	0.0363***	4.678	0.0573***	4.487
Human resources in S& T	0.0000939	0.616	-0.000302	-1.333	-0.000201	-1.144	-0.000136	-0.613
Public R& D expenditures	0.0617	0.939	0.00271	0.029	0.171**	2.814	0.254*	2.323
Private R& D expenditures	-0.00180	-0.063	-0.0366	-0.791	0.0121	0.501	-0.0182	-0.369
Relative Impact	0.129***	5.252	0.178***	5.034	0.0399**	3.110	0.0269**	2.754
Publication Share by field	0.815***	4.138	0.652	1.628	0.604***	4.370	1.833***	3.626
Wx								
GNP per capita	0.000807	0.632	0.00495*	2.338	0.00205	1.552	0.00221	1.142
Number Fields Core Membership	0.0145	0.936	-0.00455	-0.169	0.0184	1.637	-0.0137	-0.547
Human resources in S& T	-0.0000171	-0.153	0.000171	0.784	0.000293*	2.067	0.000329	1.655
Public R& D expenditures	-0.170	-1.840	-0.192	-1.374	-0.134	-1.410	-0.121	-0.771
Private R& D expenditures	0.0255	0.622	0.0929	1.092	0.0414	0.857	0.147*	2.171
Relative Impact	-0.0762*	-2.406	-0.0896	-1.682	0.0513	1.843	0.00824	0.384
Publication Share by field	-0.476	-1.703	0.768*	2.081	0.0254	0.107	0.717	0.659
Spatial								
rho	0.0535	1.532	0.127**	3.056	0.193***	4.570	0.0480	1.255
Variance								
sigma2_e	0.00817***	11.042	0.0169***	8.986	0.00543***	11.006	0.0166***	9.416
Direct								
GNP per capita	0.00188*	2.328	0.00145	1.006	0.00181*	1.980	-0.000320	-0.206
Number Fields Core Membership	0.0451***	4.427	0.0802***	4.768	0.0369***	4.750	0.0564***	4.461
Human resources in S& T	0.0000963	0.631	-0.000297	-1.313	-0.000187	-1.057	-0.000129	-0.578
Public R& D expenditures	0.0580	0.900	-0.00386	-0.042	0.166**	2.853	0.248*	2.271
Private R& D expenditures	-0.00125	-0.043	-0.0334	-0.705	0.0142	0.573	-0.0159	-0.311
Relative Impact	0.127***	5.202	0.175***	4.914	0.0421**	3.286	0.0266**	2.736
Publication Share by field	0.800***	4.153	0.659	1.645	0.602***	4.497	1.815***	3.603
Indirect								
GNP per capita	0.000863	0.635	0.00554*	2.410	0.00275	1.820	0.00213	1.046
Number Fields Core Membership	0.0181	1.091	0.00700	0.235	0.0306*	2.332	-0.0109	-0.423
Human resources in S& T	-0.0000141	-0.125	0.000148	0.618	0.000309	1.874	0.000336	1.653
Public R& D expenditures	-0.174	-1.763	-0.211	-1.335	-0.121	-1.062	-0.111	-0.671
Private R& D expenditures	0.0272	0.655	0.0995	1.083	0.0533	0.940	0.152*	2.206
Relative Impact	-0.0721*	-2.203	-0.0753	-1.243	0.0703*	2.143	0.0104	0.460
Publication Share by field	-0.465	-1.620	0.943*	2.223	0.169	0.616	0.799	0.734
Total								
GNP per capita	0.00274*	2.524	0.00699***	3.327	0.00456***	3.403	0.00181	1.205
Number Fields Core Membership	0.0632**	3.100	0.0872*	2.349	0.0675***	4.072	0.0456	1.608
Human resources in S& T	0.0000821	0.443	-0.000149	-0.488	0.000123	0.497	0.000207	0.836
Public R& D expenditures	-0.116	-0.959	-0.215	-1.176	0.0448	0.358	0.137	0.743
Private R& D expenditures	0.0260	0.534	0.0661	0.687	0.0675	1.070	0.137	1.667
Relative Impact	0.0550	1.491	0.0994	1.333	0.112**	3.183	0.0370	1.503
Publication Share by field	0.336	1.039	1.602**	2.613	0.770**	2.694	2.614*	2.218
Observations	1060		1060		1060		1060	

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

Being surrounding by rich regions (i.e. regions with higher GDP per capita) contributes to being a central region. That result confirms the significance of economically based spillover effects for centrality in scientific network (Wanzenböck et al., 2014). Such spatial dependencies may for example arise due to the location of universities, which often tend to be located in regions surrounding urban centres, with further consequence on cross-regional commuting flows between regions.

Given the total impact estimates for our regional characteristics in Medicine, the pattern observed for direct effects partly has been reversed. Total impact estimates for Medicine model show only significant positive effects for a region's GDP and private R&D expenditures. Positive capacities based on effects due to the level of public R&D expenditures, being core in other field and activities and visibility in Medicine are removed when region-internal and external impacts are considered together in our regional arrangement. These results suggest the importance of being in a rich region or surrounding by rich regions where firms do R&D. Does it mean that as Medicine is more applied science than Fundamental Biology, the presence of firms contributes to the position in the network? Moreover, Medicine seems to be more independent relative to other disciplines, as the factor "being core" is not significant.

For network centrality in Applied Biology (third column Table 5a), the economic level of the region, the effect of being core in other fields and the activities and visibility positively influence the possibility of belonging to the core. Direct impact estimates suggest that the regional production in the field is a significant important factor as well.

Positive spatial spillovers impacts on a region's network centrality can be observed for the estimates for GDP and Human Resources in Science and Technology. Thus network centrality seems to be related to the economic strength in neighbouring regions. The total impact confirms the importance of the economic level of the region and of the surrounding regions. Moreover, the number of disciplines in which the region is core is an important factor. However, the most important impact is the production of publications in Applied Biology. Quantity of scientific output matters.

For network centrality in Chemistry (last column of Table 5a), the level of publications has a positive impact. More surprisingly, relative impact of the region variable is not significant. That means that the quality of the publications does not contribute to increase the number of degrees of the region in the Chemistry network. It is the only field with such a result. The number of the fields in which the region is already core and the level of public R&D expenditures affect positively the fact of being in the core. However, even if the impact is weak, the level of Human Resources in Science and Technology has a negative impact and that seems quite puzzling.

A positive spatial spillovers impact can be observed for the estimates of the number of core fields in the surrounding regions. By contrast, the level of publications in Chemistry has a strong negative impact. That supposes that the surrounding regions must be leader in other fields but not too productive in chemistry. Negative effects may, for instance, result from core-periphery relations between neighbouring regions (Wanzenböck et al.,2014).

The total impact reduces this negative spillovers effect since only the number of core fields is significant. In Chemistry, the production of publications and their visibility would not be so important. By contrast, being core in different fields and being surrounding by regions being core in different fields largely contribute to increase the region's centrality. This is verified only for the discipline and we are inclined to interpret that as Chemistry would be more *transdisciplinary* and for that it would greatly benefit from other disciplines competence and results.

For network centrality in Physics (first column Table 5b), we observe similar magnitude and significance between our impact estimates and our coefficient estimates (available on request) confirming the minor relevance of feedback effects in publication activity between nearby regions in Physics (Wanzenböck et al., 2014). At the regional level, the field's quantity of production and quality largely contribute to the centrality of the region membership in this discipline. The experience of being core in other fields and also the level of GDP influence the centrality as well too. The only spillovers effect being significant has a negative impact and it is the relative impact of the surrounding regions. At the total effect, only being core in other disciplines and the GDP per capita are still significant. The direct and indirect effects for both the quantity and the quality in physics cancel each other out.

In the Science of the Universe (second column Table 5b), the experience of being core in other discipline and above all the relative impact of the region influence positively the centrality of the region. A positive spatial spillovers impact can be observed for the estimates of the GDP and the scientific production in the field. Being surrounding by rich regions which produce many publications in the same discipline contribute to become central.

The total impact confirms the importance of the economic and scientific level of the surrounding regions. Moreover, the number of disciplines in which the region is core is important. However, the relative impact is not anymore significant. The direct and indirect effects cancel each other out. That means that to be embedded in the science of the universe a region need to be surrounding by regions with high level of scientific production but low relative impact.

For the Engineering Sciences network (third column Table 5b), almost all the estimates are significant. That means that financial (GDP and Public R&D expenditures) and scientific resources (being core, production and quality of the research in engineering sciences) are the region-internal drivers. In particular, the scientific production and the public R&D expenditures have a strong impact.

These impacts are reinforced by the scientific production and the experience in being core in the surrounding regions. By contrast, public R&D expenditures have negative spillovers effects even if it is not significant. This one is strong enough to cancel the direct impact of the public R&D expenditure at the total impacts. To sum up, in Engineering Science, a region can become central in the co-publication network if it invests in quantity. Moreover, factors such as being core or the relative impact reinforce each other at the intern and external levels. The production and the quality of the research in Engineering Sciences build the position in the network.

For network in Mathematics (last column Table 5b), the resources dedicated to public R&D influence the centrality of the regions. The main impact is largely due to the level of scientific production in the discipline. We surprisingly observed a positive spillovers impact of the surrounding private R&D expenditures. Anyway at the total impact, only and largely accounts the level of scientific production. The relative impact is not anymore significant as both direct and indirect impacts balance each other. To sum up, to become core in mathematics, a region only needs to invest in the production of scientific publications.

Concluding remarks

In this paper, we have identified a core-periphery structure characterizing eight different scientific disciplines over the first decade of the second millennium. Independently of the discipline, a core of central regions highly connected between each other cohabites with a set of peripheral regions that try to establish connections with the most central ones. Even if different disciplines share the same network structure, much disciplinary specificity emerges. The network differs in terms of density of connections, minimum number of degree centrality necessary to be in the core and, finally, in the size of the core. However, even if these differences persist over times, degree centrality and size of the network show the same increasing pattern over time for the eight disciplines. These first results have strong methodology implications. First, heterogeneity among discipline should be taken into account: analysis of scientific network without distinguished disciplines could hide very different situation. Otherwise, specificities of discipline such as Physcs risk overshadowing the characteristics of less collaboration-intensive discipline as, for instance, Mathematics. Second, in order to consider these differences and having comparable results some normalisation method should be defined. Identifying a minimum degree of centrality for belonging to the core is the solution we propose in order to benchmark the centrality of each region.

Once this normalisation has been implemented, it is possible, how we do, to investigate the determinants of normalised centrality, i.e. being a core member. To sum up, we observe that some factors always impact the centrality of a region in the network whatever the discipline considered. Thus, at the direct impact, the experience of being core, the level of scientific production and its quality influence strongly the centrality of the region. However, we observe two exceptions: for Chemistry, the relative impact does not matter, and for the Science of the Universe the production does not have any influence. Overall, the global and the discipline scientific production and its quality are the main drivers to become member of the core. Thus, a region which aims at becoming core in one discipline has to increase its scientific production and quality. This confirms the rationality of EU policy in order to foster excellence on the one hand and to accumulate critical mass of resources on the other hand.

These direct impacts are modified slightly by the indirect impacts which are very different according to the discipline observed. This result confirms the interest of analysing the centrality discipline by discipline rather than globally and to take into account spatial effects, which, as we have seen, can be positive or detrimental.

Finally, the total impact shows that for almost every discipline, the importance of being an affluent region (GDP per capita) and with already a strong reputation in other disciplines (i.e. core in other disciplines) play a major role for determining who can enter the club of the core regions. A reinforcing mechanism is at work and this makes think that it would be difficult that core-periphery structure can be overcame in the future.

References

Almeida A., Figueiredo A., Rui Silva M., 2011, From concept to policy: building regional innovation systems in follower regions, European Planning Studies 19 (7), 1331-1356.

ASHEIM B.T. (ed.), 2006, Constructing regional advantage: principles, perspectives, policies. European Commission Report, Brussels.

European commision2010, La contribution de la politique régionale a une croissante intelligente dans le cadre de la stratégie "Europe 2020", Communication de la commission europ\'eenne COM (2010) 553 final.

COOKE P., 2001, Regional innovation systems, clusters, and the knowledge economy, Industrial and Corporate Change 10 (4), 945-974.

Doloreux D. & Dionne D., 2008, Is regional innovation system development possible in peripheral regions? Some evidence from the case of la pocatière, Entrepreneurship \& regional development, 259-283.

EUROPEAN COMMISSION, 2012, Regional innovation scoreboard, Bruxelles.

Harris R.I., Mcadam R., Reid R., 2005, A comparative analysis of innovation strategy and implementation in the UK: the effects of pheripherality, International Journal of Innovation Management 9 (4), 431-450.

Hewitt-Dundas, Roper S., 2011, Creative advantage in peripheral regions : the role of publicly funded R&D centres, Research Policy 40, 832-841

Hollanders H., Derbyshire J., Lewney R., Tijssen R., Tarantola S. and Rivera Leon L., 2012, Regional Innovation Scoreboard 2012. Methodology report. June.

OECD, 2011, Regions and innovation policy, OECD of review of regional policies, OECD.

Opsahl, T., Agneessens, F., Skvoretz, J., 2010, Node centrality in weighted networks: Generalizing degree and shortest paths, Social Networks 32, 245-251.

OST, 2010, Indicateurs de sciences et de technologies, Economica, Paris.

PINTO H., 2009, The diversity of innovation in the European Union: Mapping latent dimensions and regional profiles, European planning studies 17 (2), 303-326.

Serrano, M. A. (2008). Rich-club vs rich-multiporalisation phenomena in weighted networks, http://arXiv:0802.3122v2

TIJSSEN R.W., VAN LEEUWEN T.N., 2007, Research cooperation within Europe: bibliometric views of geographical trends and integration processes, in Torres-Salinas, D. Moed H. (eds), Proceeding trends and integration, proceedings of the ISSI 2007, CINDOC CSIC Madrid, -744

Wanzenbock I., Scherngell T., Lata R., 2012, Embeddedness of regions in European knowledge networks: a comparative analysis of inter-regional R&D collaborations, co-patents and co-publications , Annals of regional science, 337-368

WINTHJES R. AND HOLLANDERS H., 2010, The regional impact of technological change in 2020, Synthesis report for EUropean COmmission, ETEPS project.

Zelnio, R., 2012, Identifying the global core-periphery structure of science, Scientometrics 91, 601-615.

Zhou, S. and Mondragon, R.J., 2003, The Rich-Club Phenomenon in the Internet Topology, IEEE Communications Letters.

ZITT M., BASSECOULARRD E., OKUBO Y., 2000, Shadows of the past in international cooperation: collaboration profiles of the top five producers of science, Scientometrics 47 (3), 627-657.