Groupe de Recherche en
Économie Théorique et Appliquée

# Centrality of regions in R&D networks: Conceptual clarifications and a new measure

**Laurent Bergé**
*GREThA, CNRS, UMR 5113*
*Université de Bordeaux*
*Laurent.berge@u-bordeaux.fr*
*&*

**Iris Wanzenböck**
*AIT Austrian Institute of Technology, Vienna*
*and Vienna University of Economics and Business (WU),*
*Institute for Economic Geography and GIScience*
*&*

**Thomas Scherngell**
*AIT Austrian Institute of Technology, Vienna*
*and Vienna University of Economics and Business (WU),*
*Institute for Economic Geography and GIScience*

***Cahiers du GREThA***
**n° 2015-31**
**october**

**GRETHA UMR CNRS 5113**
Université de Bordeaux
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

## Centralité des régions dans les réseaux de collaboration en R&D : Clarifications conceptuelles et proposition d'une nouvelle mesure

**Résumé**

*Cet article propose une nouvelle mesure de centralité régionale dans le contexte de réseau de collaboration en R&D. Cette article discute d'abord comment les mesures de centralité de réseau existantes peuvent s'appliquer au contexte de réseau inter-régional de R&D. Par là, il démontre l'inaptitude de ces mesures à s'appliquer à ce contexte d'une façon qui a du sens. Ensuite une nouvelle mesure plus appropriée est introduite, se basant sur des connexions indirectes entre agents au niveau inter-regional. Cette mesure peut s'exprimer à partir de trois composants très simples : la participation de la région au réseau de R&D inter-régional, la part d'ouverture de la région, et la diversification de ses collaborations entre ses partenaires. On illustre ensuite la mesure et son comportement au regard d'autres mesures existantes en utilisant le réseau de co-invention Européen au niveau NUTS2.*

**Mots-clés :** centralité régionale, réseau inter-régional de R&D, réseau agrégé, réseau de co-invention

## Centrality of regions in R&D networks: Conceptual clarifications and a new measure

**Abstract**

*This paper aims at introducing a novel measure of regional centrality in the context of R&D networks. We first demonstrate some substantial problems of SNA-based centrality measures to cope with regional R&D networks in a meaningful way. Then, we introduce a new measurement approach of regional network centrality based on the concept of inter-regional bridging paths (indirect connections at the regional level). We show that the formal definition of our regional bridging centrality measure can be expressed in terms of three simple components: the participation intensity of a region in inter-regional R&D collaborations, the relative outward orientation in terms of all established links and the diversification of R&D collaborations among partner regions. We illustrate the measure and its behaviour with respect to other conventional centrality measures by using the European co-patent network at the NUTS 2 level.*

**Keywords:** network centrality of regions, inter-regional R&D networks, inter-regional bridges, aggregated networks, co-patent network

**JEL: D85, L14, O31**

**Reference to this paper: Bergé, L, Wanzenböck, I and Scherngell, T** (2015) **Centrality of regions in R&D networks: Conceptual clarifications and a new measure**, *Cahiers du GREThA*, n°2015-31.

http://ideas.repec.org/p/grt/wpegrt/2015-31.html.

**GRETHA UMR CNRS 5113**
Université de Bordeaux
Avenue Léon Duguit - 33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75 - Fax : +33 (0)5.56.84.86.47 - www.gretha.fr

# 1 Introduction

Today it is widely recognized that external knowledge sources have become an essential component for innovating organisations. Both theoretical and empirical literature over the past decade provide evidence for the increasing importance of R&D networks for successful innovation (see, e.g., Powell and Grodal, 2005; Wuchty et al., 2007). Up to now, most studies have emphasized the crucial role of the ability to adopt external knowledge in form of learning capabilities, such as technical or methodological skills, enabling innovating organisations to apply the externally tapped knowledge in the organisational innovation process. However, recently also the importance of a particular relative network positioning to access external knowledge has been highlighted and attracted increasing attention (see, e.g., Ahuja, 2000; Owen-Smith and Powell, 2004). It is assumed that not only the ability to learn, but also a favourable position for a more efficient access to external knowledge is crucial.

From a network theoretical perspective, such a favourable positioning is referred to as centrality of network vertices (Borgatti, 2005), where – in terms of R&D – these vertices represent knowledge producing actors interlinked via edges representing knowledge flows. Actors showing a more central network position will more likely benefit from network advantages. This argument has been taken up at the regional level in recent regional science literature, where regions – constituting the aggregate of its knowledge producing organisations – are treated as relevant unit of observation. In this context, the notion of inter-regional R&D collaboration networks has come into use (see, e.g., Autant-Bernard et al., 2007) where regions are the network nodes representing distinct pools of knowledge, which are assumed to get into motion via the R&D relations between these regions, constituting the edges in the network. Such a network representation has developed to an analytical vehicle that has been applied to investigate the geography of R&D networks (Scherngell, 2013), in particular how knowledge diffuses in a multi-regional system (see, e.g., Maggioni et al., 2007; Ponds et al., 2010).

Given this recent focus on regional R&D networks, network analytic measures have been increasingly applied at the regional level in order to characterize the inter-regional connectedness and centrality of a region, by capturing also the structural properties of the network (see, e.g., Sebestyén and Varga, 2013; Wanzenböck et al., 2015). For observing a region's centrality, up to now the most common analytical approaches from Social Network Analysis (SNA) have been utilized, such as degree centrality or betweenness centrality (Wanzenböck et al., 2014). However, these studies somehow neglect conceptual problems that arise for networks defined at the aggregate level of regions. In particular, such problems are related with the loss of information regarding the structure of network relations and with that, information on the real channels through which knowledge flows. In this context, the question of how to adequately reflect regions in weighted network structures such as R&D networks become even more important.

As we argue in this paper, the specific characteristics of regions – regarded as aggregate units – have to be taken into account and reflected in some way when designing analytical measurement approaches for regional centrality. Relevant questions in this context are (i) how can we conceive the centrality of regions in a network that is composed of several research actors in its underlying structure, and (ii) what are then the main building blocks that might characterize the centrality of regions, in particular when we talk about R&D networks?

This paper is one of the first that deals explicitly with the drawbacks and insufficiencies related with conventional approaches to represent networks and measure centrality at the level of regions. Against this background, the objective is to propose a new measurement approach of regional centrality that is explicitly designed for aggregated networks at the regional level, based on the concept of inter-regional *bridging paths*. Here a bridging path is defined as an indirect connection between two regions via a third 'bridging region'. From a simple random matching process that models the collaborations among the micro-level actors based on the information provided at the aggregated level, we derive a closed form of the expected number of bridges between two regions stemming from a specific bridging region. On this basis we are able to define a new measure of regional centrality that not only depends on the number of links one region has, but also on the structure and intensity of its cross-regional collaborations.

In its fundamentals, our measure of *regional bridging centrality* builds upon several network- and knowledge-related arguments, referring to the role of bridges and the relevance of bridging path between network actors, or the general importance of diversified knowledge sourcing and technological recombinations (see, e.g., Kogut and Zander, 1992; Fleming, 2001; Singh, 2005). Moreover, we show how such a measure defined for aggregated networks can be meaningfully related to the regional dimension. We demonstrate how our measure of bridging centrality of a region can be easily interpreted as a function of (i) the participation intensity of a region in inter-regional R&D collaborations, (ii) the relative outward orientation in terms of all established network links, and (iii) the diversification of network partner regions and knowledge relations to them. Hence, it views network centrality as a multidimensional problem, and integrates different region-specific aspects of the regional linking structure that might only together determine the visibility and importance of regions in R&D networks.

To illustrate our regional centrality measure we use a large-scale dataset on the European co-patent network in the year 2006 at the NUTS 2 level. The comparative analysis with three common SNA-based measures (degree, betweenness and eigenvector centrality) is based on basic statistics on distribution and correlations between the four centrality measures observed for the regional network. Despite striking similarities in correlations and distributional aspects on a more general level, the in-depth analysis of regional ranks reveals interesting differences which emphasize the advantages of the regional bridging centrality measure, in particular in terms of its interpretative power for region-level analyses.

The remainder of this study is structured as follows: Section 2 discusses in some detail the conventional approach to measure the centrality of regions in R&D networks. Section 3 introduces the concept of bridging paths, constituting the main essence of the measurement approach proposed in this study, before Section 4 formally derives the bridging centrality measure for regions. Section 5 shifts attention to the illustrative example, applying our measure to the European co-patent network and comparing results with conventional measures, before Section 6 concludes with a summary of the main results and some ideas for future research.

# 2 The conventional measurement approach

The notion of the centrality of regions in regional R&D networks has come into use just recently. It is argued that the knowledge creation ability within a region depends to a large extent on the ability of the region-specifc actors to efficiently access region-external knowledge (see, e.g., Bathelt et al., 2004; Graf, 2011). Inter-regional R&D networks are regarded as effective means in this regard with network links representing direct channels to a specific (region-external) source of knowledge that actors otherwise would not have access to. Against this background, need has been expressed to derive analytical approaches to measure a region's centrality in such networks, enabling the empirical researcher to characterize whether a region has a favourable position in the network, whether it takes a specific – for instance 'brokering' – role from a global network perspective, or how a region's network positioning changes over time.

However, the concept of network centrality was originally defined at the individual level in human communication networks and the implications of using this concept at the regional level remain unclear. Therefore, this section intends to clarify the concept of network centrality as applied to inter-regional knowledge networks. We start with examining the origin, meaning, and purpose of network centrality (Subsection 2.1), and then lay out the major hurdles facing its transposition to regional R&D networks (Subsection 2.2). Finally, the two last subsections provide different ways to adapt well known centrality measures to the regional case, while at the same time keeping focus on their interpretation in the R&D context and pointing out their conceptual limitations.

## 2.1 A short introduction to the notion and context of centrality measures in social networks

The inception of the use of the concept of centrality in social network analysis (SNA) lies on the impetus of Bavelas' early researches (Bavelas, 1948). He was interested in linking the relational position of individuals within working-groups – namely their network-centrality – to their performance and influence over the group. Many empirical studies have followed to investigate if such a link existed in these type of networks, i.e., human communication networks (e.g., Bavelas, 1950; Leavitt, 1951; Faucheux and Moscovici, 1960; Burgess, 1969). The consequence of this line of work was to unveil the potential of the concept of network centrality in SNA.

As the representation of interactions in a network-form is not limited to human communication networks, the notion of centrality was soon extended and applied to various other types of networks. Indeed, this idea of investigating the influence of structural position within networks was promising and has triggered many researches where the unit of analysis took different forms. Such studies include the application of the notion of centrality on: inter-personal networks within organizations (Beauchamp, 1965), cities in transportation networks (Pitts, 1965), the diffusion of innovation in inter-firm informal communication networks (Czepiel, 1974), the spread of diseases in infection networks (Bell et al., 1999), crime networks (Calvó-Armengol and Zenou, 2004), etc.

Along with these studies, a set of centrality measures has also emerged. Indeed, numerous measures have spawned either to refine existing measures or to adapt them to the networks under scrutiny. Those centrality measures include: the degree centrality, the betweenness (Freeman, 1977), the closeness (Freeman, 1979), the eigenvector (Bonacich, 1972), Katz's prestige (Katz, 1953), Bonacich's measure of power (Bonacich, 1987), etc.

Consequently, as a wide variety of centrality measures has been developed, one should expect that they differ in the meaning they purport and in the contexts they can be applied to. These differences are in fact tied to the very definition of network centrality.

The goal of a centrality measure is to assign to each agent of a network a value related to her/his position within the network. The variety of centrality measures then comes from the fact that each favours a particular network-pattern over others and each carries a 'view' of what being central *should be*. Thus, centrality measures are not neutral: they rank the agents along some – often hidden – normative viewpoint which should support the aim of the study itself. In other words, different notions of centrality imply different 'competing "theories" of how centrality may affect group process' (Freeman, 1979, p. 238).

Then, the choice of a centrality measure should be dictated by the purpose it is aimed to serve (Borgatti, 2005). This purpose is brought about by the researcher and his research study and is of course highly context dependent. For instance, the kind of centrality measure used in the study of infection networks should be different from the one used in inter-firm cooperation networks.[1] This very idea is also, albeit slightly differently, formulated by Bonacich (1987,

[1] In the study of spreading disease in infection networks, the notion of eigenvector centrality catches best the idea that the central agent, if infected, would spread the fastest the disease across the network (Borgatti, 1995). When studying flows of information in inter-firm communication networks, the closeness centrality

p. 1181):

> There are different types of centrality, depending on the degrees to which local and global structures should be weighted in a particular study and whether that weight should be positive or negative. [...] There is no point in subsuming all these situations under one measure.

Therefore, there is no unique and 'best' measure of centrality, no 'one size fits all' centrality measure. One then should remember the implicit choices underlying centrality measures and the context to which they can be applied. Therefore, we are now going to discuss the particular context of regional R&D networks and question whether centrality measures can be applied to it.

## 2.2 Can the concept of network-centrality be applied to R&D networks?

We now delineate two key elements impeding the straightforward application of network-centrality measures to regional R&D networks. *First*, regions are not single entities. Indeed, while being at the centre of the analysis, regions are not the 'actors' taking part to the action of the network. Only the agents that compose the regions are involved in R&D networks (and any kind of inter-regional network more generally). Centrality measures are best suited for situations where the unit of analysis is also the actor of the network. In fact, in the case of regional centrality, there is a strong duality between the micro strata, where lie the actors of the network, and the meso strata, where lies the focus of the centrality measure. Indeed, to assimilate regions as 'actors' would imply to assume that all agents within them would act as one and only one entity; it would require to do 'as if' the region was a single agent, like for instance a single researcher. If this 'as if' hypothesis may be reliable when studying small groups in which information is quickly shared and without depreciation, such as research teams or even – under some conditions – organizations, it no longer holds when looking at complex structures such as regions which are often composed of heterogeneous, non necessarily interacting, agents.[2]

*Second*, the links in R&D networks involve a particular kind of flows. For instance, in collaboration networks, a link may be the medium of various types of exchanges and could then be interpreted in different ways. If we focus specifically on the notion of knowledge production, the links can represent the access to a specific source of knowledge that agents would have otherwise not have access to, like the possibility to share tacit knowledge with a

reports the best the idea that the central agent would be the first to 'know' the novelties and by then have a technological edge over its competitors (Czepiel, 1974; Freeman, 1979).

[2]It is to note that Everett and Borgatti (1999) propose an extension of centrality measures to groups but where within-group homogeneity is required to provide a proper interpretation.

partner (Collins, 2001). If the focus is more on the dynamics of the collaboration network, links can be seen as vehicles of information, on who would be a suitable and a reliable partner to collaborate with, particularly across regional borders (see, e.g., Gulati and Gargiulo, 1999; Cassi and Plunket, 2015). These two simple different perspectives on how to interpret network links have different implications. In the first case, in which we consider flows of knowledge, the benefits from network-distant agents may decay much more steeply than for the case of flows of information which is acquired and shared more easily. These differences in flows' nature and behaviour are not innocuous regarding the interpretation of centrality measures, as Borgatti (2005, p. 69) has pointed out: 'the importance of a node in a network cannot be determined without reference to how traffic flows through the network'. He has also shown that different centrality measures each carry an implicit different assumption about the kind of flow it is suited for, so that they cannot be applied to any network.

With these details in mind, the next subsection considers the case in which regional R&D networks are seen as weighted networks. Some widely used centrality measures are described as well as: 1) their classic interpretation in the context in which they were originally defined and 2) their interpretation when applied to regional R&D networks. Finally, last subsection investigates the case in which a region's centrality is inferred by its agents centralities.

## 2.3 Regional R&D networks as weighted networks

The first manner to adapt existing centrality measures to regional R&D networks is to consider the regions as the nodes of the network. Accordingly, the inter-regional R&D collaboration network can be depicted by the matrix $G$ of typical element $g_{ij}$ which represents the number of links between the agents from regions $i$ and $j$. As collaborations are bilateral and their flow can be higher than one, it yields an undirected weighted matrix $G$ of typical element $g_{ij} \in \mathbb{R}^+$.

We discuss three conventional measurement in this case: the degree-, the eigenvector- and the betweenness-centrality. The properties of these centrality measures are discussed in light of the context of R&D networks.

The first centrality measure, is the degree-centrality. The notion of degree-centrality in SNA was primarily defined as the number of connections an agent had in communications networks and is reviewed in Freeman (1979). As Freeman mentions, early researchers in SNA even considered it as the sole centrality measure, able to summarize the importance of a node in a network. In the case of regional networks, the links between two nodes are typically weighted, the degree of a node can then be defined as the sum of all the links stemming from it.[3] Let $d_i$ be the degree of node $i$, it is formally defined as follows: $d_i = \sum_j g_{ij}$.[4] Depending on the kind of network under study, the degree can be interpreted as the probability to be reached in a network by a random walk or the ability to infect other agents in a one time period (Borgatti, 2005). However, these two interpretations hardly make sense in the case of R&D networks. Another simple and unambiguous interpretation of the degree is just the dominance of a given region over other regions in terms of R&D collaborations. Depending on the purpose of the study, this interpretation may be relevant. However, in any case, this measure suffers from a major flaw: it does not convey any information on the structure of the network.

Another centrality measure widely used in SNA is the eigenvector centrality. This measure was introduced by Bonacich (1972) and states that the importance of a node is related to the importance of the nodes it is connected to. Contrary to the degree-centrality, the eigenvector-centrality of a given node depends on the information on all the links of the network, meaning the position of the nodes within the global network has an influence on their centrality. Therefore, two nodes with the same degree can have different eigenvector centralities. Formally, the eigenvector-centrality of a node, $e_i$, is defined by the relation: $\lambda e_i = \sum_j g_{ij} e_j$, with $\lambda > 0$ a proportionality factor. This centrality is self-referential and can be solved by writing it in a matrix-form:

$$\lambda \mathbf{e} = G\mathbf{e}, \tag{1}$$

where $\mathbf{e}$ is the vector of all centralities. The vector $\mathbf{e}$ that solves equation (1) is the eigenvector of the matrix $G$ associated to the eigenvalue $\lambda$.[5] The very idea reflected by this measure is related to node influence. The main driver is that a node will be more influential if it has influence on very influential nodes (the influence being measured by the links between the nodes). While being an appealing feature for studies on individual's influence, this interpretation is strongly impeded by the problem of the micro/meso duality of the regional network. Indeed, assume a region is central thanks to connections to important regions, do its agents – who are the actors of the network – really benefit from their region's centrality? It would imply that every agent within a region would homogeneously benefit from the influence of all other agents of the region, which seems hardly the case. It then happens that this measure is hardly transposable to the regional level.

[3] In the case where the network is directed, like for instance in a patent-citations network, the number of links emanating from a node (e.g., references made to other patents) is called the out-degree while the number of links received (e.g., the number of citations received from other patents) is called in-degree. For undirected networks, such as collaboration networks, the in-degree is equal to the out-degree.

[4] There is a generalization of the degree centrality for weighted networks given by Opsahl et al. (2010) but whose interpretation in this context remains unclear.

[5] By convention, it is standard to use the eigenvector associated to the largest eigenvalue (Bonacich, 1987; Jackson, 2010).

A third measure commonly applied in SNA is the betweenness-centrality. To define it, we first introduce the notion of network path and shortest path. A path between two nodes $i$ and $j$ is a sequence of $K$ distinct nodes $\{n_1, \ldots, n_K\}$ starting from $i$ (i.e., $n_1 = i$), ending with $j$ (i.e., $n_K = j$) and such that each consecutive pair of nodes is connected in the network.[6] The length of a path is the number of nodes composing the path. Then, a shortest path between $i$ and $j$ is a path that has minimal length. Now, let $SP(jk)$ to be the number of shortest paths between nodes $j$ and $k$, and $SP_i(jk)$ the number of shortest paths between $j$ and $k$ where node $i$ appears. Then the betweenness-centrality of $i$ is defined by the following equation:

$$B_i = \sum_{j \neq i} \sum_{k \neq \{i,j\}} \frac{SP_i(jk)}{SP(jk)},$$

the term in the double sum depicting the share of shortest paths between $j$ and $k$ where $i$ lies on.

This form of centrality was originally defined in the context of communication networks, where links between agents represent information flows. When Freeman introduced this measure, he defined central agents as 'structurally central to the degree that they stand between others and can therefore facilitate, impede or bias the transmission of messages' (Freeman, 1977, p. 36). Alternatively, betweenness-centrality can be seen as how much a node is necessary for flows to connect all other nodes in the network. Despite being computationally easy to apply at the regional level, this measure suffers from major flaws when applied to regional R&D networks. Indeed, for the importance of being in the 'shortest path' to hold, two assumptions are necessary. The first is that the flows necessarily follow the shortest path (which makes the 'central agent' able to retain information and exert some influence). If information (or the adequate flow) does not pass only through shortest paths, this measure becomes much less relevant. In R&D networks, this may not be the case: for instance, when considering information over potential partners obtained via collaboration, that information may not be limited to flow only through shortest paths, just because of the nature of information. The second assumption is that flows do not suffer from any decay. Indeed, at the moment where the relevance of network-flows are reduced with the network-distance, then what is the use of being in the middle of network-paths between agents? In this case, the betweenness of a region, beyond its first or second circle of connections, may be of little use. For instance, if connections materialize access to knowledge sources, it is quite unlikely that agents far apart with respect to the network-distance influence each other. Last, beyond these two limiting assumptions, the betweenness measure happens to be much better suited for networks composed of individuals (be it firms or inventors). As for the eigenvector centrality, its interpretation hardly fits the regional scale as a region with a high

[6] Mathematically, $\{n_1, \ldots, n_K\}$ is a path between $i$ and $j$ if $g_{n_k n_{k+1}} > 0$ for all $k \in \{1, \ldots, K-1\}$, with $n_1 = i$ and $n_K = j$.

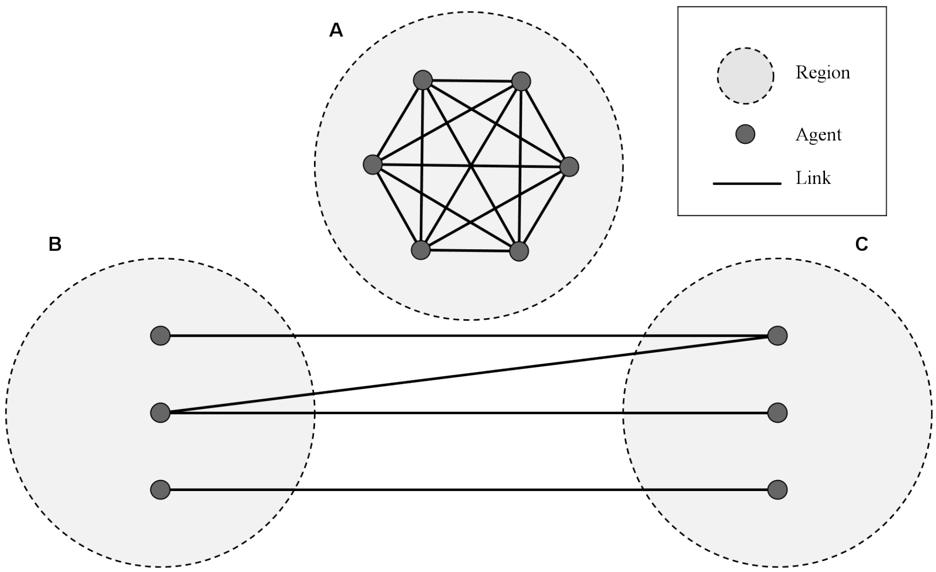


Figure 1: Illustration of a regional network in which a region has a strong internal structure yet no link with the outside.

betweenness does not necessarily translates into its agents being on shortest paths.

As we have shown, existing measures can be applied to regional R&D networks, when taking regions as the nodes of a weighted network. But the interpretation of the measures and their conceptual meaning is far from being straightforward, if applicable at all. In the next section we show and discuss another way of accounting for regional centrality.

## 2.4 Regions as the aggregate centrality of their actors

A different way to measure regional centrality is to assume that a region's centrality actually refers to the centrality of its agents. Indeed, since regional networks can be seen as the aggregate interactions of the agents from these regions, a natural way to assess a region's centrality could be to link it to the centrality of its agents.

In doing so, the first step is to find the relevant actor of the network. In co-patenting networks, it can either be firms or inventors. The choice depends on whether we believe that the information and knowledge pool of firms is shared among all its inventors. If so, then firms can be considered as the real actors of the network. We will call 'agent' the entity resulting of this choice. Thus, the regional network can then be depicted by a micro-level network formed of the links between the agents, each of them belonging to a region. To build the regional centrality, one has to choose the relevant centrality measure and compute it at the agent's level. Let $c_i$ be the centrality of agent $i$ and let $S_r$ be the set of agents belonging to region $r$. Then, the centrality of region $r$, $C_r$, can be defined merely as the sum of the

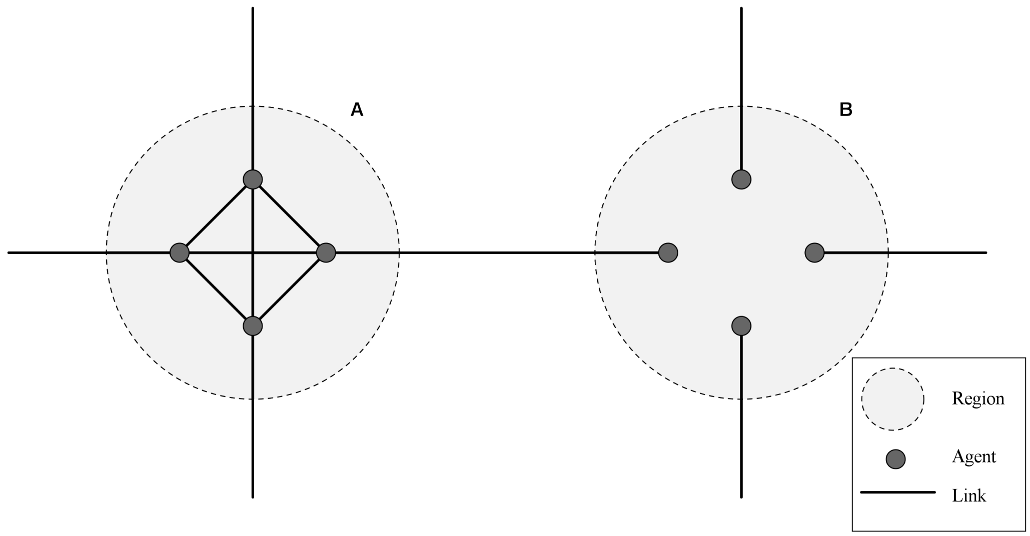


Figure 2: Sample of an inter-regional network. Illustration of two regions with external links, differentiated with respect to their internal links.

centrality of its agents, as follows:

$$C_r = \sum_{i \in S_r} c_i.$$

Beyond the problems inherent to the centrality measures (e.g., linked to the nature of the flows), discussed earlier, this methodology – aggregation of micro-level network centralities – also involves drawbacks. The main problem stems from the links occurring internally to regions. Indeed, should intra-regional links be counted in the micro-level network? An example of a problematic case is illustrated by Figure 1. In this figure, a network of three regions is represented: region A has many agents that are all connected to each other but have not any link with other regions; conversely, the agents from regions B and C have no intra-regional link but do have cross-regional collaborations. Actually, if the network is computed at the micro-level, whatever the measure, region A will have the higher centrality, despite having no inter-regional link whatsoever. This is fundamentally problematic: a measure of regional centrality should not be able to rate high regions having no external links simply because it should somewhat relate to the position within the interregional network which is not the case here.

A straightforward solution to this problem would be to 'cut' all intra-regional links: the centrality would then be computed using a network where all internal links would be severed. Yet, this adjustment would also lead to conceptual problems. Take for instance the example illustrated by Figure 2. This figure depicts a network of two regions, A and B, that are very similar. They both are composed of four agents and each have a link with an other region. Region A has a strong internal structure: all its agents are connected. On the contrary, no agent from region B does have any link within the region. Despite that the agents of the two regions have different positions in the global network, if all internal links are cut to compute the centralities, then the two regions would be equivalent. Cutting internal links would involve a distortion in the network structure.

Consequently, the major problem raised by the aggregation of micro-level centralities is that intra-regional links cannot either be kept or removed without posing conceptual problems.

As developed in this section, the centrality measures discussed so far all suffer from conceptual drawbacks when applied at the regional level. Given these considerations, there is a need for developing alternative centrality measures applicable for regional R&D networks and resting on more robust conceptual grounds. In what follows, we provide a first attempt for the development of novel measurement approaches that explicitly address the conceptual problems discussed above by taking into account the underlying micro structure of regional R&D networks.

# 3 The concept of bridging paths

There is a strong need for overcoming the duality in analyses of R&D networks of regions concerning the micro level which encompasses the actors participating in R&D collaborations, and the aggregate, i.e. regional, level where the analysis focuses on. As has been discussed in the previous section, major problems arise in applying and interpreting conventional SNA-based centrality measures. The purpose of this section is to provide a new concept that is *meaningful* in the context of inter-regional R&D networks. We introduce the notion of a bridging path denoting a form of indirect connection between regions, i.e. regions are indirectly connected in the network thanks to their micro-level actors. We first define this concept before providing an approach to derive the expected number of bridging paths from aggregate flows of R&D interactions. The expected number of bridging paths between regions will be the major building block of the regional centrality measure we introduce in the next section.

To introduce the concept of bridging paths, consider a network where the nodes are the regions and the connections between the regions represent the R&D interactions between their agents. This represents a weighted network where we define $g_{ij}$ as the number of R&D interactions (i.e. micro-level links) between regions $i$ and $j$. Further, each micro-level link between two regions is denoted by $y^a_{ij}$, where $y^a_{ij}$ represents the $a^{th}$ link between regions $i$ and $j$ with $a \in \{1, \dots, g_{ij}\}$. A bridging path is then regarded as a set of two links at the micro level connecting three agents from three different regions. Speaking in social network analytical terms, the micro-level agent in one region act as a 'broker' (Burt, 1992) for two other not directly connected actors; he/she has a bridging role in the network of regions linking indirectly the micro-level agents of two other regions. This triangulation between actors located in three different regions leads to the notion of an inter-regional bridging path. Formally, a bridging path is defined as a set of two links from two different regions, say $i$ and $j$, with a third one, say $k$, so that the agents from $i$ and $j$ are both connected to the same agent in $k$. This means that a pair of links $(y^a_{ik}, y^b_{jk})$ forms a bridging path if, and only if, $y^a_{ik}$ and $y^b_{jk}$ are connected to the same agent in region $k$. In other words, agents $i$ and $j$ are indirectly connected thanks to one agent of region $k$.

This notion is depicted by figure 1 which represents a regional network of three regions. In this figure, the pair of links $(y^2_{ik}, y^1_{jk})$ is a bridging path between regions $i$ and $j$ stemming from $k$ because the agent from $k$ maintains both links $y^2_{ik}$ and $y^1_{jk}$. Although both regions $j$ and $k$ do have links with region $i$, there is no bridging path between them because the agents from $i$ of the links $y^1_{ik}$ and $y^2_{ik}$ are neither connected to $y^1_{ij}$, $y^2_{ij}$ nor $y^3_{ij}$. Hence, region $i$ provide not any bridging path between regions $j$ and $k$ in this set-up. We see that the notion of bridging path is about indirect connections. Accordingly, the region with most bridging paths is region $j$, as it provides two bridging paths between regions $i$ and $k$.

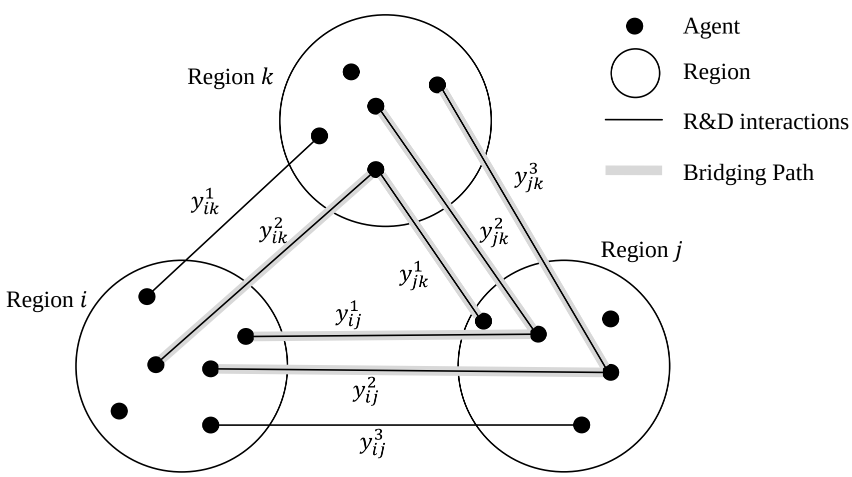


Figure 3: Illustration of the notion of bridging paths

*Notes*: The figure depicts three bridging paths formed by the following pairs of links: $(y^2_{ik}, y^1_{jk})$, $(y^1_{ij}, y^2_{jk})$ and $(y^2_{ij}, y^3_{jk})$. So the regional dyads $(j,k)$, $(i,k)$ and $(i,j)$ have respectively 0, 2 and 1 bridging paths stemming from regions $i$, $j$ and $k$, respectively.

The relevance of the bridging paths concept can be quite directly underlined by means of basic theoretical considerations in innovation research. The creation of new knowledge is often viewed as a recombination of existing knowledge (see, e.g., Kogut and Zander, 1992; Fleming, 2001; Cassiman and Veugelers, 2006). It implies that the source from which the agents draw their knowledge will have an impact on their ability to generate interesting ideas and new knowledge. In the case where a region is isolated, where collaborations occur mainly within the region, the knowledge pool may become redundant and even lead to lock-in situations (see, e.g., David, 1985; Arthur, 1989). Collaborations with agents from other regions allow to benefit from different knowledge bases (see, e.g., Singh, 2005; Berliant and Fujita, 2012), help moderate the problem of redundancy and generate more radical innovations. From this viewpoint, bridging paths provide better knowledge opportunities to regions. Then regions from which stem many bridging paths could be seen as key players in the network with actors potentially benefiting from a more diversified knowledge pool.

Moreover, bridging paths may also be of significance when we consider network formation processes. Indeed, several recent studies have put at the forefront the consideration that the structure of network links plays an important role in explaining future states of the network (see, e.g., Barabâsi et al., 2002; Jackson and Rogers, 2007). Recent research in the context of R&D networks has shown that two actors are more likely to collaborate together if they share a common collaborator (that is if they are indirectly linked in the network, see, e.g., Fafchamps et al., 2010; ter Wal, 2014). Hence, bridging path create network proximity and opportunity for (triadic) closure so that there are good reasons to assume that bridging paths matter for the evolution of the whole network. Indeed, if bridging paths represent indirect connections between agents from different regions, then we can assume that those regions which provide the bridging paths are in a position to facilitate the connectivity between other regions in the network. Bridging paths can then be seen as important for regions not only in the context of accessing a diversified knowledge pool, but also in a network formation perspective as it helps the formation of inter-regional connections and with that inter-regional diffusion of knowledge.

# 4 A new measure of regional centrality

Proposing the significance of the bridging path concept for measuring regional centrality in regional R&D networks, the question arises at this point how this concept can be incorporated into regional centrality measures. Usually, empirical researchers focusing on regions as units of observations, and by this, on regional R&D networks, face the problem that the underlying micro structure of the network may be either undefined or unobservable. Concerning the latter, one may consider the example of co-patenting networks (see, e.g., Lata et al., 2015), for which the relevant actors are individual persons (inventors) that are hardly identifiable as homogeneous nodes over time. Thus, we introduce a model of random matching. It allows us to approximate the underlying micro-structure by deriving an expected number of bridging paths (ENB) between two regions, using only the aggregate flows of collaborations between regions.[7]

Our random matching process relies on two basic assumptions: (i) collaborations occur between two agents, and (ii) when a collaboration occurs, the two agents are matched at random. By this, it reflects the ex post probability to be matched, i.e. the probability that two agents for two particular regions have been matched conditional to the structure of the inter-regional flows of collaborations. The very intention is to give a baseline for a micro-network that was likely to occur, with respect to what is observable at the meso level. Thus, random matching is used to infer the structure of the micro network by using only the information included the links between the regions.

On this basis, it is now possible to derive the expected number of bridging paths stemming from a given region by using directly the aggregate flows of collaborations occurring between regions. First, denote by $n_i$ the number of actors active in R&D collaboration in region $i$. Then the expected number of bridging paths, $ENB^i_{jk}$, between the two regions $j$ and $k$ stemming from the bridging region $i$ along the random matching process is:[8]

$$ENB^i_{jk} = \frac{g_{ij} g_{ik}}{n_i}. \tag{2}$$

The expression related by equation (2) simply states that the more connections two regions, $j$ and $k$, have with a third common region, $i$, the more likely they will have indirect connections at the micro level (bridging paths) thanks to the actors located in $i$.

Based on this, we are able to construct a new measure of the centrality of regions in R&D networks, denoted as *regional bridging centrality (BC)*. The BC is defined as the number of bridging paths stemming from a region between all dyads of the network. Formally, this means that the BC of region $i$ is equal to:

$$BC_i = \sum_{j \neq i} \sum_{k \neq i,j} ENB^i_{jk}, \tag{3}$$

where $ENB^i_{jk}$ is defined by equation (2).

The interesting point of our measure is that its definition can be pretty much simplified and interpreted meaningfully in a regional context. Assume that the number of agents ($n_i$) is proportional to the number of projects ($g_i$); then equation (3) decomposes to a notion of centrality of a region that entails a combination of three different components, reflecting i) a region's *participation intensity,* ii) a region's *relative outward orientation* and iii) a region's

[7] This model is an adaptation of the one in Bergé (2015). In fact, the methodology is very similar to the one used by Bloom et al. (2013), which provides a measure of technological similarity between firms' patenting activity introducing a model which considers random encounters between pairs of scientists.

[8] For a formal proof, see Bergé (2015).

*diversification of network links*.[9] It is defined as

$$BC_i \;=\; \bar{g}_i\, s_i\, (1 - h_i)\,, \tag{4}$$

where

$\bar{g}_i$ is the number of outer collaborations (i.e. outer degree, that is $\bar{g}_i = g_i - g_{ii}$ which is the total number of collaborations of $i$, noted $g_i$, excluding the internal ones, noted $g_{ii}$). It refers to a region's *participation intensity* in inter-regional collaborations, which affects positively the centrality of the region. It is a general measure of how well a region is embedded in the particular R&D network. Note that a region's size will amplify the probability of yielding more bridges between other regions. The participation intensity could therefore be interpreted as a broad measure of the relational capacity of the regional network nodes, which should be taken into account.

$s_i$ is the share of outer collaborations with $s_i = \bar{g}_i / g_i$. It can be related to the *relative outward orientation* of all established network linkages, i.e. the relative degree of external R&D interactions. It refers to the openness of a region with respect to knowledge sourcing strategies. Given the fact that the BC focuses on the capacity of one region to link other regions, a high number of region-internal collaborations would have a negative influence as it potentially reduces the number of actors connecting different regions.

$h_i$ refers to the Herfindahl-Hischman (HH) index of the distribution of $i$'s outer collaborations defined as $h_i = \sum_{j \neq i} (g_{ij} / \bar{g}_i)^2$ . The term $1 - h_i$ varies between 0 and 1 according to the degree of *diversification of network links* to other regions, and indicates how a region's collaborations are distributed along its neighboring regions in the network. In this case, the more the collaborations are concentrated, the less the region is central. This is because concentration offsets the benefits of outer connections as it reduces the actors' possibility to build bridges among different regions. Also it relates to the fact that the more the outer collaboration pool is diversified over different regions, the more the region can draw its knowledge from different sources.

One central promising property of the measure is that it takes account of the peculiar characteristics of regional networks. Indeed, regional networks are characterised by the structure of region-internal and region-external links and this feature cannot be dealt with adequately by using a single (a-spatial) SNA centrality measure. A region's ability to benefit from new ties in the R&D network or exploit external knowledge sources via the links may be determined

[9] The formal proof is given in Appendix A.

by all three components together. Outward orientation and higher diversification in particular may help a region to develop and renew the regional knowledge base faster, or prevent lock-in situations in certain technologies (see, e.g., Breschi and Lenzi, 2015).

# 5 An illustrative example: an application to the European co-patent network

Given the promising features of the regional bridging centrality (BC) measure as defined in the previous section, an application to empirical regional R&D networks is required in order to illustrate the behaviour of the measure as compared to the conventional ones. To this end, we will employ co-patent data, comparing the regional BC with three other commonly used centrality measures, that is degree, eigenvector and betweenness centrality.[10] We use the European co-patent network, a network of inter- and intra-regional collaborations in patent production observed at the regional level. A co-patent, that is a collaboration issuing a patent grant, is a visible trail of a successful R&D collaboration and is defined as an invention implying at least two inventors. This data are extracted from the REGPAT database (Maraut et al., 2008) and consist of all patents applied for at the European patent office (EPO) in the year 2006. We make use of the information contained in each patent record to build the co-patent network. Particularly, we use the address contained in each inventor's byline to map every patent to a set of NUTS 2 regions. That is, the NUTS 2 regions represent the place of residence of the inventors when they applied the patent. We consider that the flow of inter-regional collaborations between two regions consists of all patents having at least one inventor from each of these two regions. Collaborations occurring strictly within the regions are counted as intra-regional patent.

The network consists of collaboration flows between 245 NUTS 2 regions. This cross-regional co-patenting network is based on a total of 40,142 patents, of which 16,661 are inter-regional collaborations linking the 245 NUTS 2 regions. As a starting point, the three components of the BC are described by table 1a. The participation intensity is on average 237, which means that the regions show on average 237 co-patent links to other regions in the network. This is much higher than the median of 100, confirming the right-skewed distribution of the number of co-patent links the individual regions hold to other regions.

More interestingly is the relative outward orientation. Here, the median is 71%, meaning that for half the regions, more than 71% of their patents are of inter-regional nature, being

[10]The degree is here calculated as the number of unique projects the agents of a region are involved in. The eigenvector and the betweenness centrality are computed using the package `igraph` available in the statistical software `R`. Both these two measures are based on the weighted regional co-patent network where the nodes are the regions and where the linkages between any two regions are the number of patents co-invented by agents from these two regions. Due to the nature of the network, we used the weighted version of both the betweenness and the eigenvector centrality.

invented with at least one partner outside the regions. Also diversification is relatively high, with an average at 85%, meaning that the co-patents are rather distributed along several regions. Hence, the regions resort – on average – to a rich portfolio of partner regions leading to a diversified structure of inter-regional knowledge exchanges in patenting. In contrast to the participation intensity, the other two components, the relative outward orientation and the structure, are slightly left skewed, and can be seen as moderators of the scale of a region. Indeed, being a large region with a high network participation intensity does not necessarily lead to a high centrality value, if either the share of intra-regional collaborations is very large or inter-regional links are concentrated among only a few regions.

Table 1 reports some statistics on the BC measure as compared to the conventional measures, and the correlations among them. Note that all measures are normalized so that the highest value is one and the lowest zero.[11] While there is no large difference in the summary statistics provided by table 1b, it can still be noted that the eigenvector-centrality is clearly the more skewed of the four measures. Table 1c further shows that the correlation between the bridging centrality and the other measures ranges from 70% to 93%. Those high levels are reassuring as they show that the BC does not completely reorder the regional positioning. The difference in the distribution of the four centrality measures compared is also illustrated by figure 2 which reports the cumulative distribution of each measure. We can see that the eigenvector-centrality, except at the very beginning of the distribution, is on the top of all other measures while the BC lies between the degree and the betweenness. The differences in distribution are higher at the beginning of the distribution (below 0.50) than at the end where the distribution of the BC, the degree and the betweenness are much closer. Yet, the differences with existing measurements are real and it is worthwhile to point out to changes occurring to some particular regions. Moreover, it becomes obvious from this basic statistics that the bridging centrality is a combination of three components. It depends not only the scale of a region, like it might be the case for the degree centrality, or the quality of partners, i.e. whether they are located at the very core of the network, as for the eigenvector centrality. Therefore, it might be of particular interest how differently the three components are distributed across the individual regions.

[11] Formally, the transformation applied to each centrality measure is: $(x - x_{min})/(x_{max} - x_{min})$.

Table 1: Descriptive statistics of the components of the BC and of the centrality measures applied on co-patenting data.

(a) Descriptive statistics of the three components of the bridging centrality measure.

| | Min | Q1 | Median | Q3 | 90% | Max | Mean | SD | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| Participation intensity | 1 | 26 | 100 | 280 | 559 | 2333 | 237.16 | 376.77 | 3.09 | 10.78 |
| Relative outward orientation | 0.2000 | 0.600 | 0.737 | 0.835 | 0.907 | 1 | 0.714 | 0.16 | -0.45 | -0.14 |
| Diversification | 0 | 0.831 | 0.893 | 0.925 | 0.945 | 0.972 | 0.850 | 0.14 | -3.85 | 18.47 |

(b) Summary statistics.

| | Min | Q1 | Median | Q3 | 90% | Max | Mean | SD | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| Bridging Centrality | 0.0000 | 0.0083 | 0.0332 | 0.0965 | 0.2231 | 1.0000 | 0.0881 | 0.1490 | 3.3465 | 13.2502 |
| Degree | 0.0000 | 0.0116 | 0.0406 | 0.1384 | 0.2597 | 1.0000 | 0.1081 | 0.1722 | 3.0310 | 10.1794 |
| Eigenvector | 0.0000 | 0.0010 | 0.0035 | 0.0208 | 0.0927 | 1.0000 | 0.0434 | 0.1285 | 4.9103 | 26.8212 |
| Betweenness | 0.0000 | 0.0022 | 0.0118 | 0.0595 | 0.1719 | 1.0000 | 0.0598 | 0.1307 | 4.3528 | 23.0673 |

(c) Correlations.

| | Bridging Centrality | Degree | Eigenvector | Betweenness |
|---|---|---|---|---|
| Bridging Centrality | 1.0000 | 0.9080 | 0.9311 | 0.6886 |
| Degree | 0.9080 | 1.0000 | 0.8141 | 0.8411 |
| Eigenvector | 0.9311 | 0.8141 | 1.0000 | 0.5641 |
| Betweenness | 0.6886 | 0.8411 | 0.5641 | 1.0000 |

Notes: The *participation intensity* is the outer degree. The *relative outward orientation* is the share of outside collaborations over all collaborations, it varies between 0 and 1. The *diversification* is $1-h_i$ with $h_i$ being the Herfindahl index of the distributions of region $i$'s collaborations over all other regions; it varies between 0 and 1, the more the collaborations are concentrated, the lower is the measure. 9

Table 2: Centralities of the top 30 regions for the co-patent network, ranked by bridging centrality.

| | NUTS 2 | Bridging Centrality value (rank) | Degree Centrality value (rank) | Eigenvector Centrality value (rank) | Betweenness Centrality value (rank) |
|---|---|---|---|---|---|
| Karlsruhe | DE12 | 1.00 ( 1) | 0.90 ( 3) | 1.00 ( 1) | 0.46 ( 7) |
| Darmstadt | DE71 | 0.85 ( 2) | 0.86 ( 5) | 0.79 ( 3) | 0.53 ( 5) |
| Rheinhessen-Pfalz | DEB3 | 0.80 ( 3) | 0.66 ( 8) | 0.89 ( 2) | 0.25 (12) |
| Düsseldorf | DEA1 | 0.80 ( 4) | 0.82 ( 6) | 0.62 ( 4) | 0.51 ( 6) |
| Köln | DEA2 | 0.78 ( 5) | 0.73 ( 7) | 0.59 ( 6) | 0.55 ( 4) |
| Oberbayern | DE21 | 0.57 ( 6) | 0.93 ( 2) | 0.39 ( 7) | 0.93 ( 2) |
| Stuttgart | DE11 | 0.51 ( 7) | 0.87 ( 4) | 0.59 ( 5) | 0.34 ( 8) |
| Northwestern Switzerland | CH03 | 0.50 ( 8) | 0.44 (13) | 0.18 (16) | 0.24 (14) |
| Freiburg | DE13 | 0.48 ( 9) | 0.55 ( 9) | 0.35 ( 9) | 0.20 (20) |
| Arnsberg | DEA5 | 0.42 (10) | 0.39 (17) | 0.31 (10) | 0.06 (62) |
| Berlin | DE30 | 0.40 (11) | 0.42 (14) | 0.22 (13) | 0.18 (22) |
| Tübingen | DE14 | 0.38 (12) | 0.47 (12) | 0.35 ( 8) | 0.15 (31) |
| Île de France | FR10 | 0.34 (13) | 1.00 ( 1) | 0.06 (36) | 1.00 ( 1) |
| Münster | DEA3 | 0.33 (14) | 0.29 (19) | 0.22 (12) | 0.16 (28) |
| Mittelfranken | DE25 | 0.33 (15) | 0.40 (16) | 0.16 (18) | 0.11 (37) |
| Alsace | FR42 | 0.30 (16) | 0.26 (26) | 0.13 (19) | 0.22 (16) |
| Zurich | CH04 | 0.30 (17) | 0.32 (18) | 0.11 (22) | 0.18 (23) |
| Schwaben | DE27 | 0.29 (18) | 0.28 (20) | 0.21 (14) | 0.06 (58) |
| Brandenburg | DE40 | 0.28 (19) | 0.23 (29) | 0.16 (17) | 0.03 (92) |
| Hannover | DE92 | 0.25 (20) | 0.25 (27) | 0.12 (21) | 0.08 (50) |
| Unterfranken | DE26 | 0.24 (21) | 0.26 (24) | 0.23 (11) | 0.05 (64) |
| Rhône-Alpes | FR71 | 0.24 (22) | 0.54 (10) | 0.06 (37) | 0.32 (10) |
| Hamburg | DE60 | 0.24 (23) | 0.21 (35) | 0.10 (25) | 0.09 (47) |
| Prov. Vlaams-Brabant | BE24 | 0.23 (24) | 0.19 (38) | 0.04 (44) | 0.13 (33) |
| Espace Mittelland | CH02 | 0.22 (25) | 0.26 (25) | 0.09 (26) | 0.06 (60) |
| Koblenz | DEB1 | 0.22 (26) | 0.18 (40) | 0.19 (15) | 0.07 (57) |
| Schleswig-Holstein | DEF0 | 0.22 (27) | 0.22 (32) | 0.10 (23) | 0.04 (72) |
| Prov. Antwerpen | BE21 | 0.21 (28) | 0.21 (34) | 0.04 (42) | 0.19 (21) |
| Lüneburg | DE93 | 0.20 (29) | 0.17 (43) | 0.07 (34) | 0.10 (44) |
| Région de Bruxelles, Brussels Hoofdstede | BE10 | 0.20 (30) | 0.14 (61) | 0.02 (56) | 0.06 (59) |

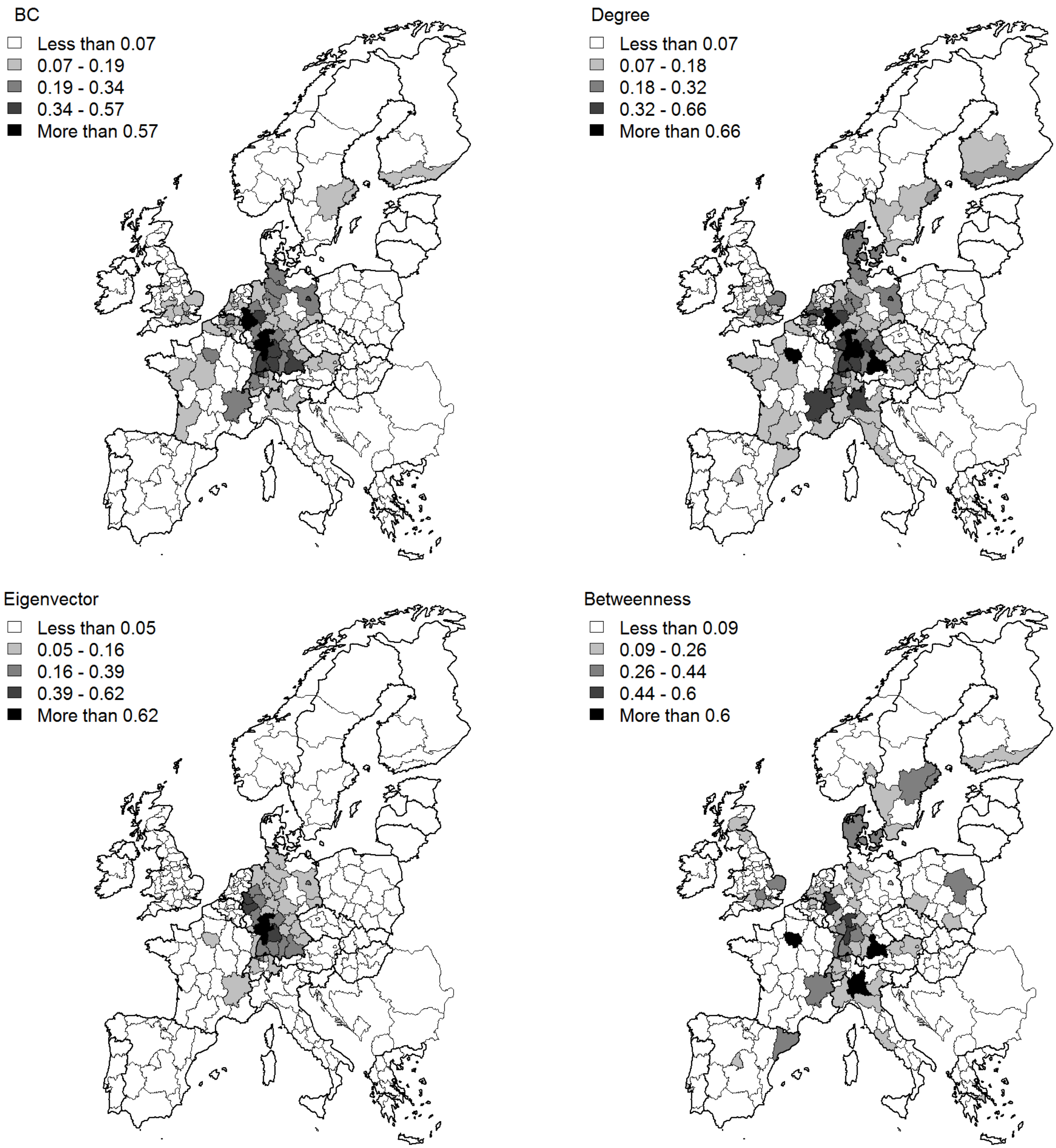


Figure 4: Spatial distribution of the four centrality measures among the 242 NUTS 2 regions.

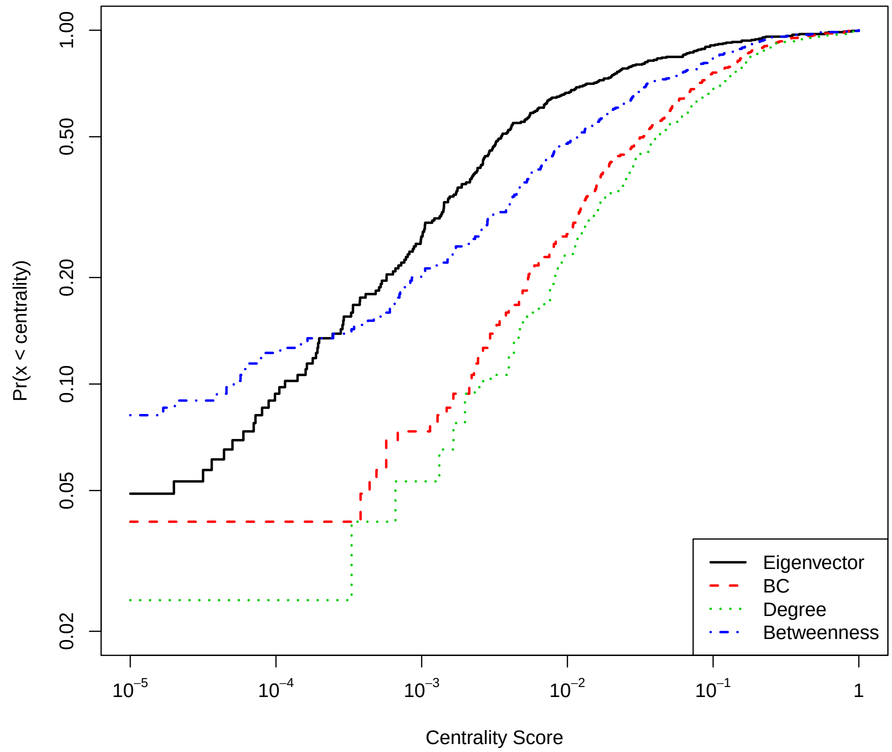


Figure 5: Cumulative distributions of the centrality measures in log-log.

Table 2 represents the top 30 centralities ordered by the bridging centrality. We focus on commenting the most salient differences. As highlighted by Figure 4, the ranking is clearly dominated by German regions which rank highest for most measures. Interestingly, we find 13 German regions among the 15 best ranked regions for the bridging centrality. This results from the fact that they show both a high participation intensity as well as high openness from an inter-regional perspective; they show a high absolute as well as relative number of inter-regional co-patents. However, the concentration tendency and high clustering of co-patenting activities at the national level of Germany may point to the fact that economic linkages at the national level prevail. Likely explanations are low language / cultural barriers as well as lower transaction costs. These factors seem to promote the high regional bridging centrality in German regions.

Another interesting case is the region of Île de France (FR10) which ranks at the $13^{th}$ position for the bridging centrality, while being ranked first with respect to its degree centrality. We see that the measure of degree centrality may overstate its position in the inter-regional co-patent network. Despite its highly distributed structure of collaborations (it has a low HH index of 0.04), this region is highly reliant on internal collaborations (the outer share of collaborations is only 45%) that it fails to provide much bridging paths to the inter-regional R&D network. By contrast, the eigenvector centrality may understate the importance of

FR10; it ranks only 36 as it is linked to the network core regions at a lower degree. For the same reason as for FR10, some regions that are ranked high in the degree centrality end up much lower in the BC; i.e. they show high embeddedness in the inter-regional R&D network but are less open and diversified in the structure of their inter-regional collaboration, thus receiving lower values of bridging centrality.

Following the criteria of openness and diversification, interesting is also the case of Brussels (BE10) which ranks after the $56^{th}$ place for all centrality measures other than the bridging centrality. With the BC, BE10 ranks $30^{th}$, gaining at least 26 places compared to other measures. Yet, the SNA-based centrality measures may underestimate its positioning in the inter-regional co-patent network: due to its very high outward orientation (its outer share is 94%) and a highly distributed structure of collaborations (it has a low HH index of 0.07), this region is likely to provide many bridging paths to the network and may therefore be an important bridge for the whole network and for inter-regional knowledge diffusion.

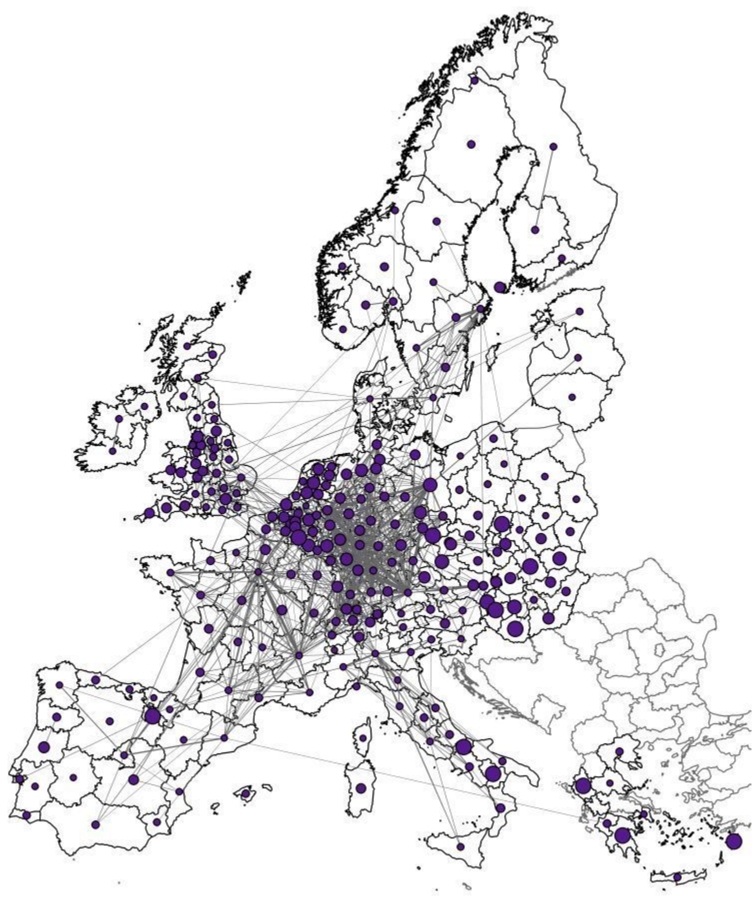

Figure 6: The European co-patent network

*Notes:* Node size corresponds to the relative outward orientation of a region, line width corresponds to the number of co-patents between two region.

Figure 3 illustrates the European co-patent network for the European NUTS 2 regions, with the node size corresponding to the relative outward orientation of a region. It confirms the very dense network structure between core regions clustered in Germany, which hold intensive connections among each other. From a regional perspective, the bridging centrality is high for these regions, i.e. they yield high values for all three components, despite the fact that most of the links are confined at the national level. Furthermore, we observe a high relative outward orientation of some South and Eastern European regions. In terms of established co-patent links they seem to be highly open, which could be explained by the lack of internal collaboration structures. Nevertheless, inter-regional linkages are generally weak for these regions.

# 6 Concluding remarks

The notion of centrality is ubiquitous in debates on the role of regions in R&D networks. Quantitative approaches to measure regional centrality, however, are often based on micro-level centrality measures as introduced in social network analysis (SNA). Empirical analysis of regional networks requires accounting for the network structure originally defined at the micro level or by the linkages between different organisations, which often limits the usefulness and conclusive identification of regions in the network. A further unavoidable problem relates to the considerable loss of information regarding network structure and meaning when regions are regarded only as aggregate units. In this study we address this micro / meso-level duality in how we view regional networks and the region's structural network positioning is usually defined, questioning the conventional measurement approaches for region-level analysis.

By introducing the notion of regional bridging centrality we suggest a new approach for assessing the centrality of regions in R&D networks that is able to cope with the regional dimension in measuring the centrality. Based on the concept of bridging paths, i.e. a set of two links connecting three actors in three different regions, we develop a measure of centrality that satisfies the requirements of both R&D networks and region-level applications: A bridging path between regions characterizes a situation where regional actors represent bridges or brokers in the network of regions as they connect indirectly the actors located in two other regions. Such a triangulation in regional networks, as we argue, is a key issue for knowledge recombinations and the extension of a region's knowledge base.

We further show that centrality in terms of bridging centrality can be viewed as a function of (i) the participation intensity in inter-regional collaborations, (ii) its openness to other regions (i.e. the relative outward orientation of network links), and iii) the diversification of links to other regions. With these three components – which are both intuitive and computationally simple – we argue that regional network centrality has to be viewed from a multidimensional perspective. Only with such an integrative perspective we can achieve a better understanding of the role of certain regions in inter-regional R&D networks.

The comparative analysis with three standard SNA-centrality measures confirms the performance and usefulness of our measure of regional bridging centrality. We chose the inter-regional co-patent network for European regions as illustrative example. Despite observing similar patterns in basic statistics like correlations of the centralities or the skewness, we were able to show striking and interesting differences in the structure of the inter-regional co-patent linkages across regions. The results reveal that thinking only of the degree of participation is not enough. Rather, the most central regions show simultaneously high embeddedness, high relative outward orientation and high diversification of their network links (e.g. Karlsruhe). In contrast, regions that may be strongly embedded (i.e. high participation intensity) may show low openness and diversification of links, thus yielding lower centrality values (e.g. Île de France). Hence, a region's outward orientation and the diversification of its network links moderates the influence of regional scale on network centrality. This is a major strength of the measure proposed in this study, and it paves the way for future studies to examine the role of certain regions in networks of inter-regional knowledge flows. Viewing network positioning of regions in terms of regional bridging centrality might further elevate our understanding of which regions are the most central, show high visibility and at the same time are most important for the network and the inter-regional diffusion of knowledge.

Furthermore, the bridging centrality measure may contribute to the development of a multi-dimensional typology of regions, based on structural network criteria according to their levels of embeddedness, openness and diversification of links in inter-regional networks. Such a typology might enhance our understanding of how different the roles of regions in networks might be, and how they contribute to the arrangement and evolution of the inter-regional structure. This is one of our main points for a future research agenda. Moreover, it seems natural that an application of the bridging centrality measure on other types of knowledge networks according to different technological fields might reveal interesting patterns of the most central network nodes. Hence, the measure of bridging centrality is not limited to the context of R&D collaborations but may prove to be useful also for the application in other types of network structures, such as inter-regional trade flows or inter-regional economic value chains, also regarding their evolution over time.

# References

Ahuja, G., 2000. Collaboration networks, structural holes and innovation: A longitudinal study. Administrative Science Quarterly 45 (3), 425–455.

Arthur, W. B., 1989. Competing technologies, increasing returns, and lock-in by historical events. Economic Journal 99, 116–131.

Autant-Bernard, C., Mairesse, J., Massard, N., 2007. Spatial knowledge diffusion through collaborative networks. Papers in Regional Science 86 (3), 341–350.

Barabâsi, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., Vicsek, T., 2002. Evolution of the social network of scientific collaborations. Physica A 311 (3), 590–614.

Bathelt, H., Malmberg, A., Maskell, P., 2004. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. Progress in Human Geography 28 (1), 31–56.

Bavelas, A., 1948. A mathematical model for group structures. Human organization 7 (3), 16–30.

Bavelas, A., 1950. Communication patterns in task-oriented groups. The Journal of the Acoustical Society of America 22, 271–282.

Beauchamp, M. A., 1965. An improved index of centrality. Behavioral Science 10 (2), 161–163.

Bell, D. C., Atkinson, J. S., Carlson, J. W., 1999. Centrality measures for disease transmission networks. Social Networks 21 (1), 1–21.

Bergé, L. R., 2015. Network proximity in the geography of research collaboration. Papers in Evolutionary Economic Geography, Number 15.07 Working paper.

Berliant, M., Fujita, M., 2012. Culture and diversity in knowledge creation. Regional Science & Urban Economics 42 (4), 648 – 662.

Bloom, N., Schankerman, M., Van Reenen, J., 2013. Identifying technology spillovers and product market rivalry. Econometrica 81 (4), 1347–1393.

Bonacich, P., 1972. Factoring and weighting approaches to status scores and clique identification. Journal of Mathematical Sociology 2 (1), 113–120.

Bonacich, P., 1987. Power and centrality: A family of measures. American journal of sociology 92, 1170–1182.

Borgatti, S. P., 1995. Centrality and aids. Connections 18 (1), 112–114.

Borgatti, S. P., 2005. Centrality and network flow. Social Networks 27 (1), 55–71.

Breschi, S., Lenzi, C., 2015. The role of external linkages and gatekeepers for the renewal and expansion of US cities knowledge base, 1990-2004. Regional Studies 49 (5), 782–797.

Burgess, R. L., 1969. Communication networks and behavioral consequences. Human Relations 22 (2), 137–159.

Burt, R. S., 1992. Structural holes: The social structure of competition. Harvard University Press, Cambridge, MA.

Calvó-Armengol, A., Zenou, Y., 2004. Social networks and crime decisions: The role of social structure in facilitating delinquent behavior. International Economic Review 45 (3), 939–958.

Cassi, L., Plunket, A., 2015. Research collaboration in co-inventor networks: Combining closure, bridging and proximities. Regional Studies 49 (6), 936–954.

Cassiman, B., Veugelers, R., 2006. In search of complementarity in innovation strategy: Internal R&D and external knowledge acquisition. Management Science 52 (1), 68–82.

Collins, H. M., 2001. Tacit knowledge, trust and the Q of sapphire. Social Studies of Science 31 (1), 71–85.

Czepiel, J. A., 1974. Word-of-mouth processes in the diffusion of a major technological innovation. Journal of Marketing Research 11, 172–180.

David, P. A., 1985. Clio and the economics of QWERTY. American Economic Review 72 (2), 332–337.

Everett, M. G., Borgatti, S. P., 1999. The centrality of groups and classes. The Journal of Mathematical Sociology 23 (3), 181–201.

Fafchamps, M., van der Leij, M. J., Goyal, S., 2010. Matching and network effects. Journal of the European Economic Association 8 (1), 203 – 231.

Faucheux, C., Moscovici, S., 1960. Etudes sur la créativité des groupes tâche, structure des communications et réussite. Bulletin du C.E.R.P. 9, 11–22.

Fleming, L., 2001. Recombinant uncertainty in technological search. Management science 47 (1), 117–132.

Freeman, L. C., 1977. A set of measures of centrality based on betweenness. Sociometry 40 (1), 35–41.

Freeman, L. C., 1979. Centrality in social networks conceptual clarification. Social Networks 1 (3), 215–239.

Graf, H., 2011. Gatekeepers in regional networks of innovators. Cambridge Journal of Economics 35 (1), 173–198.

Gulati, R., Gargiulo, M., 1999. Where do interorganizational networks come from? American Journal of Sociology 104 (5), 1439 – 1493.

Jackson, M. O., 2010. Social and economic networks. Princeton University Press.

Jackson, M. O., Rogers, B. W., 2007. Meeting strangers and friends of friends: How random are social networks? American Economic Review 97 (3), 890–915.

Katz, L., 1953. A new status index derived from sociometric analysis. Psychometrika 18 (1), 39–43.

Kogut, B., Zander, U., 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. Organization Science 3 (3), 383–397.

Lata, R., Scherngell, T., Brenner, T., 2015. Integration processes in European research and development: A comparative spatial interaction approach using project based research and development networks, co-patent networks and co-publication networks. Geographical Analysis, 1–27.

Leavitt, H. J., 1951. Some effects of certain communication patterns on group performance. The Journal of Abnormal and Social Psychology 46 (1), 38–50.

Maggioni, M. A., Nosvelli, M., Uberti, T. E., 2007. Space versus networks in the geography of innovation: A European analysis. Papers in Regional Science 86 (3), 471 – 493.

Maraut, S., Dernis, H., Webb, C., Spiezia, V., Guellec, D., 2008. The OECD REGPAT Database. OECD Science, Technology and Industry Working Papers.

Opsahl, T., Agneessens, F., Skvoretz, J., 2010. Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks 32 (3), 245–251.

Owen-Smith, J., Powell, W. W., 2004. Knowledge networks as channels and conduits: The effects of spillovers in the Boston biotechnology community. Organization Science 15 (1), 5–21.

Pitts, F. R., 1965. A graph theoretic approach to historical geography. The Professional Geographer 17 (5), 15–20.

Ponds, R., Oort, F. v., Frenken, K., 2010. Innovation, spillovers and university-industry collaboration: an extended knowledge production function approach. Journal of Economic Geography 10 (2), 231–255.

Powell, W. W., Grodal, S., 2005. Networks of innovators. In: The Oxford Handbook of Innovation. Oxford University Press, Oxford, pp. 56–85.

Scherngell, T. (Ed.), 2013. The geography of networks and R&D collaborations. Springer-Physica Verlag, Berlin-Heidelberg-New York.

Sebestyén, T., Varga, A., 2013. Research productivity and the quality of interregional knowledge networks. Annals of Regional Science 51 (1), 155–189.

Singh, J., 2005. Collaborative networks as determinants of knowledge diffusion patterns. Management Science 51 (5), 756 – 770.

ter Wal, A. L. J., 2014. The dynamics of the inventor network in German biotechnology: geographic proximity versus triadic closure. Journal of Economic Geography 14, 589–620.

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

Wanzenböck, I., Scherngell, T., Lata, R., 2015. Embeddedness of European regions in European Union-funded research and development (R&D) networks: A spatial econometric perspective. Regional Studies 49 (10), 1685–1705.

Wuchty, S., Jones, B. F., Uzzi, B., 2007. The increasing dominance of teams in production of knowledge. Science 316 (5827), 1036–1039.

# Appendix

## A Obtaining the Bridging Centrality

Assume that the number of agents of region $i$, $n_i$, and the number of projects of that region, $g_i$, are proportional so that $n_i = \alpha g_i$. Then the bridging centrality can be rewritten as follows:

$$
\begin{aligned}
BC_i &= \sum_{j\neq i}\sum_{k\neq i,j} ENB^i_{jk} \\
&= \sum_{j\neq i}\sum_{k\neq i,j} \frac{g_{ij}g_{ik}}{\alpha g_i} \\
&= \frac{1}{\alpha}\frac{1}{g_i}\sum_{j\neq i}\left[g_{ij}\sum_{k\neq i,j} g_{ik}\right] \\
&= \frac{1}{\alpha}\frac{1}{g_i}\sum_{j\neq i} g_{ij}\left(\bar{g}_i - g_{ij}\right) \\
&= \frac{1}{\alpha}\frac{\bar{g}_i^2}{g_i} - \frac{1}{g_i}\sum_{j\neq i} g_{ij}^2 \\
&= \frac{1}{\alpha}\frac{\bar{g}_i^2}{g_i}\left(1 - \sum_{j\neq i}\left(\frac{g_{ij}}{\bar{g}_i}\right)^2\right) \\
&= \frac{1}{\alpha}\bar{g}_i s_i\left(1 - h_i\right)
\end{aligned}
$$

Further, as the $\alpha$ is common to all regions, we lose no generality to setting it to $\alpha = 1$. Which yields the result. $\square$

# *Cahiers du GREThA*
# *Working papers of GREThA*

**GREThA UMR CNRS 5113**

Université de Bordeaux

Avenue Léon Duguit
33608 PESSAC - FRANCE
Tel : +33 (0)5.56.84.25.75
Fax : +33 (0)5.56.84.86.47

**http://gretha.u-bordeaux.fr/**

## Cahiers du GREThA (derniers numéros – last issues)

*2015-19 :* MALFAIT Jean-Jacques, MARTIN Jean-Christophe, *Identification d'une filière économique liée à l'utilisation d'une ressource naturelle renouvelable : l'exemple de la filière bois en France et en Aquitaine*

*2015-20 : COMBRIS Pierre,* GIRAUD-HERAUD Éric, Alexandra Seabra PINTO, *Relative willingness to pay and surplus comparison mechanism in experimental auctions*

*2015-21 :* GIRAUD-HERAUD Éric, PONSSARD Jean-Pierre, SINCLAIR DESGAGNE Bernard, SOLER Louis-Georges, The Agro-Food Industry, Public Health and Environmental Protection: the Porter hypothesis to justify regulation ?

*2015-22 :* BOKINO Régis, Transparence dans les deux banques centrales de la zone franc africaine : évaluation et comparaison internationale

*2015-23 :* DUCRAY François, *Les aides aux étudiants - Une question d'arbitrage entre les bourses sociales et la demi-part fiscale*

*2015-24 :* FLAMAND Marina, FRIGANT Vincent, *L'hétérogénéité des comportements des constructeurs automobiles vis-à-vis du Corporate Venture Capital*

*2015-25 :* BRESCHI Stefano, LISSONI Francesco , MIGUELEZ Ernest*, Foreign inventors in the US: Testing for Diaspora and Brain Gain Effects*

*2015-26 :* ASSOGBA Guillaume, KLEBANER Samuel*, Vers un cadre d'analyse institutionnaliste de la politique de filière : Quelle cohérence pour la politique de filière française ?*

*2015-27 :* LALANNE Aurélie, ZUMPE Martin*, Zipf's law, Gibrat's law and Cointegration*

*2015-28 :* CHARLES Léo, *Evolution of trade patterns and economic performance: the case of France and Switzerland during the nineteenth century*

*2015-29 :* LLOPIS Oscar*, Who do you care about? Scientists' personality traits and perceived beneficiary impact*

*2015-30 : ZUMPE Martin,* Le nouveau modèle keynésien : une interprétation d'«Interest and Prices» de Michael Woodford

*La coordination scientifique des Cahiers du GREThA est assurée par Emmanuel PETIT.*
*La mise en page est assurée par Julie VISSAGUET*