The determinants of network formation: a dynamic perspective

Lorenzo Cassi
lorenzo.Cassi@univ-paris1.fr
Université Paris 1, CES - Centre d'Economie de la Sorbonne

And

Anne Plunket
anne.plunket@u-psud.fr
Université Paris Sud 11, ADIS

Abstract:
This paper investigates the dynamics of network formation in genomics patenting in France (1986-2007) and it tests the respective role of cumulative mechanisms (e.g. closure) and actor strategies (e.g. filling a structural hole) in determining the evolution of inventor networks. To do that, we consider and analyze four types of co-inventor links. Taking two individual inventors as our focal point, four categories of links are distinguished: (1) a link bridging two components; (2) a link determining the creation of a new component; (3) a pendant to an existing component; or (4) an intra-component link. We investigate four determinants of dyad formation: the inventors' network position, the geographic and technological proximity between inventors, and besides these more usual determinants, we propose to consider how the applicants' strategies in terms of cumulative and bridging affect the inventor's link formation.

First draft (28.02.10)

Introduction

An increasing body of literature investigates innovation networks and their dynamics. These networks play a major role for both strategic and economic reasons.

From a strategic viewpoint, actors within innovation networks seek partners in order to gain access to complementary knowledge, competences or resources. Their role and position in these networks have a number of potential consequences for their own performance and the network structure. As networks span across geographic boundaries they can also broaden the knowledge base and the competences and open the network to non-local actors.

From an economic perspective, the location of networks is of particular importance for a number of regions in terms of potential growth and development. Typically, innovation networks are strongly geographically clustered around some large agglomerations. This increases interaction frequency and, as a result, increases the number of common projects and improves their efficacy.

Since networks are crucial for innovation and actors, as well as for regions' performance, it seems crucial to consider the conditions under which these networks are formed and the relative importance of factors acting as major network drivers.

In this paper we investigate the dynamics of network formation in genomics patenting in France (1986-2007), we test the respective role of cumulative mechanisms (e.g. closure) and actor strategies (e.g. filling a structural hole) in determining the evolution of inventor networks.

1. Theoretical framework and background literature

Starting from the seminal work of Jaffe, Henderson and Traintenberg (1993), patent data have been analyzed to track knowledge flows and to show the importance of geographical distance in determining the potential beneficiaries of those externalities. However, as Breschi and Lissoni (2001) have argued, localized knowledge spillover has long been treated as a black box and only recently have some empirical works started shedding some light on their role.

Many of these recent attempts have been based on patent data and the analysis of inventor networks has become a rather rich subfield.

While over the past decade this corpus of empirical works has reached some generally accepted conclusions, and has helped establishing Social Network Analysis as key methodological tool, there is as yet no consensus on important substantial issues.

These analyses cleared highlighted the key role of social networks in determining knowledge diffusion and exchange. Major results concern in particular the role played by social distance in knowledge diffusion as well as the role of cross-firm inventors in establishing network connectivity. Relative to this latter point, the role of academic inventors has also been widely studied and emphasized.

Recently, an important methodological issue has been tackled: identifying and handling homonymy among inventors (the so-called john smith problem or name game) has been recognized as non-trivial task, with important consequences in terms of robustness of results.

A still under-investigated question concerns the dynamic of network formation. Indeed this has recently been emphasized as a major research agenda for the geographical analysis of innovative network (Boshma and Frenken, 2009).

The analytical framework proposed by Glucker (2007) appears as a suitable reference to address this issue, namely the definition of geographical network trajectory he puts forward:

> *It describes a geographically and historically specific development path of a network in which the formation and dissolution of ties in earlier stages generates cumulative properties for the formation and dissolution of ties in the future and in which the mechanisms of path-disruption and variation are endogenous (Glucker, 2007)*

According to this author, the evolution of network therefore results from two orthogonal sets of mechanisms. The first is a cumulative mechanism set, related to the historical process, where both initial conditions and the observed sequence of events matter. The second is a selective mechanism, which deals with the strategies that individual actors implement in order to gain benefit of participating in the network, given the constraints represented by the present structure of the network itself.

The former mechanisms account for cumulative mechanisms such as, for instance, the formation of cliques. The simple idea is that a cluster of individuals tends to form a clique (i.e. a group of actors where everyone is connected to everyone) in order to fulfill social obligations. An individual actor can meet some difficulties as she fails to collaborate with someone else who in turn collaborates with the majority of her collaborators. This is exactly one of the costs of redundancy mentioned by Burt, i.e. the loss of autonomy.

In more general terms, social network literature *lato sensu*, identified three different cumulative mechanisms: preferential attachment (the tendency to link to the most connected actors), homophilia (the tendency to form a link to most similar actors) and closure (the tendency to cluster formation).

Selective mechanisms, on the other hand, could be interpreted more explicitly in terms of economic rationality: actors tend to allocate their resources efficiently, given some constraints. According to this logic, actors choose to form a link with actors endowed with specific assets (e.g. relevant knowledge) or to fill structural holes in order to profit from their strategic position (e.g. control the information flow).

The goal of this paper is to identify the different determinants of these mechanisms in order to explain the overall evolution of the network.

2. Patent networks in genomics

2.1. Description of the data and network formation

The dataset under investigation is composed of all the genomic patents published at the European Patent Office, between 1986 and 2007, with at least one inventor reporting a French address. We therefore considered all the inventors of a genomic patent located in France and their co-inventors whatever their location within or outside France. A genomic patent is defined according to the IPC classification codes. These codes were identified in a recent research project carried out by ADIS-Paris Sud, INRA (Nantes) and the OST – Observatoire des Sciences et des Techniques - and supported by the French national research agency (ANR).

Our dataset includes 2400 patents, 5261 inventors and 662 applicants.

In order to build the network, standard five-year windows were considered. For instance, the network in 1994 was built up taking in account all the patents (inventors) published between 1990 and 1994. Accordingly, an inventor was considered as active (e.g. in 1994), if she had at least one patent over the 1990-1994 periods.

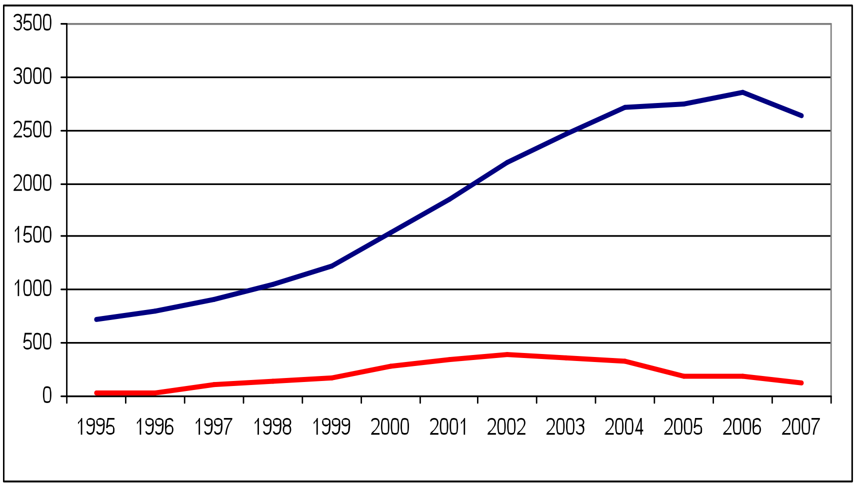
Contrary to Fleming et al. (2007), our networks are not regionally bounded but could include ties from other regions within and without France. This enabled us to consider geographical distance, as a determinant in network formation.

2.2. Networks structural properties

Figure 1 displays the number of active networks over time. At the beginning of 2000, the number of inventors clearly grows and then stabilizes around 2004.

More striking is the time-varying pattern depicted by the giant component size: first, it appears to be relatively small throughout the period by comparison to the size estimated in other similar analysis (e.g. Fleming and Frenken, 2007). Second, it reaches its maximum in the year 2002, and starts decreasing before reaching a plateau.

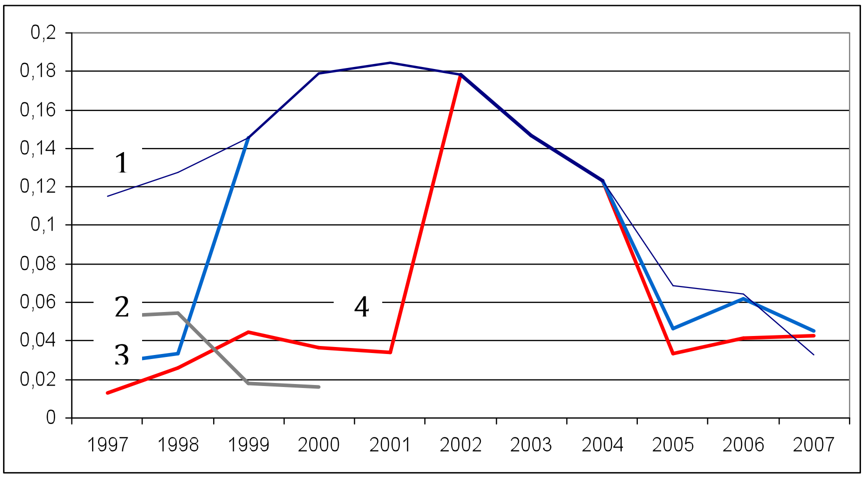
Figure 1. Active inventors (blue line) and giant component size (red line)



While previous analyses only took into account the giant component's dynamics, this paper also tracks the dynamical network aspects outside the giant component (Baum et al., 2003; Lee Fleming et al. 2007; Fleming and Frenken, 2007).

It is interesting to consider the formation of the giant component over time and understand why some network subparts become connected and grow over time whereas others do not. The formation of the largest component may be the result of two scenarios that are not necessarily mutually exclusive. In the first, the largest component may result from the connection of relatively large existing components that increase over time, have their own dynamics and finally become connected in a larger one. In the second scenario, the largest component may result from an incremental process wherein small components become connected, within a short time period, to a single relatively large component. Figure 2 reports the evolution of the first four components in the 1998 network.

Figure 2. 1998 first four components over time



The first component (137 inventors in 1998, around 13% of the active inventors) is mainly composed of inventors located in Ile-de-France (the same holds for the second and the fourth components), while the bulk of the third component is located in the Rhône-Alpes region. The components also differ in terms of patent applicants. The first component includes some big pharmaceutical corporations (e.g. Aventis and Centillion) and some foreign universities; the second mainly includes public actors as CNRS, INSERM and some Parisian universities as well as biotechnological firms (e.g. Neurotech SA). Finally, the third component revolves around one main applicant, Bio Merieux, while the fourth one is mainly composed of inventors working for a spin-off of CSIRO, the Australian government research agency. The most striking is that the 'public' component, i.e. the second one, breaks up during the last few considered years, while the other components converge to a giant component. Finally, in the most recent years, the size of the giant component drops down with its members splitting into three subgroups.

3. The determinants of networks formation

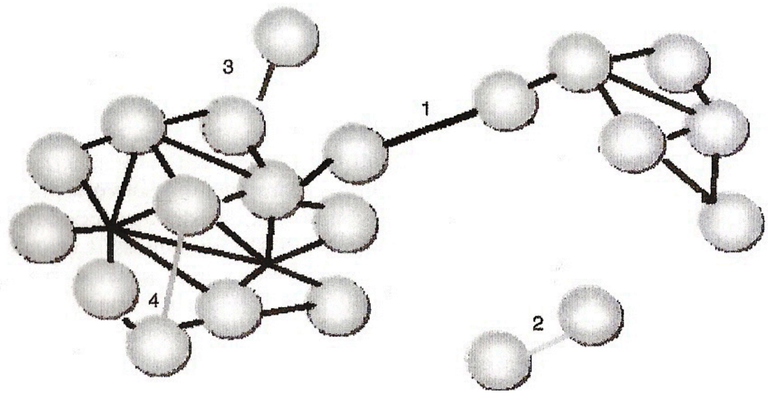
3.1 Network structure and dyads

The objective of our paper is to understand the evolution of the overall network structure through the formation of co-inventor links. For this reason, our unit of analysis is the dyad. Namely, we analyze the determinants of new tie formation between all inventors from 1991 to 2007.

Following Amburgey et al. (2008), we classified each link as represented in figure 3. Taking two individual inventors as our focal point, they may become connected through four categories of links: (1) a link bridging two components; (2) a link determining the creation of a new component; (3) a pendant to an existing component; or (4) an intra-component link. In this draft only the first (bridging) and the last (intra-component) links are considered.

The formation of each type of link has different implications in terms of the overall structure of the network. For instance, the creation of a bridging link –type 1– has direct consequences on the number of existing components (decreasing) and their size (increasing), while the formation of a new intra-component link –type 4– does not affect the number neither the number of component nor their size but their density.

Figure 3. Type of networks ties



Source: Fig. 10, Amburgey et al. (2008)

Table 1 reports the number and share of new links relative to the period 1991-2007 in our data.

Table 1. New link: type of networks ties

Links	Total number	%
1. Bridging links	265	1,79
2. New components link	10037	67,91
3. Pendant links	4266	28,87
4. Intra-components links	211	1,43
Total	14779	100

Our econometric exercise consists of analysing the determinants of each type of links. The comparison among the estimated models will permit us to highlight the different mechanisms underlying the formation of each type.

A number of factors may explain why two inventors become connected. We consider in this paper the inventors' characteristics (e.g. their network position, technological specialization and geographical location) as well as the strategy of applicants in terms of knowledge base enhancing or deepening through bridging and closure mechanisms. We do not necessarily expect that one determinant dominates than the other one. Indeed we can expect that the formation of link will depend on the interaction of individual behavior and firms' strategies, that one of these two could result to be more important in determining the formation of one type of link rather than the others. Moreover, we intend to test the dynamic properties and see if some changes in terms of the relative importance of these two determinates can occur over time.

3.2. Network effects and inventor's position

Former studies on patent data have highlighted the small worlds properties (Watts and Strogatz's, 1998) of the co-patent networks (Fleming, King III and Juda, 2007; Ter Wal, 2009). These are formed by groups of actors that appear strongly clustered with low path length that increase the probability of forming new ties with once co-inventors' former partners. On the one hand, since all inventors are rather connected within small cluster and not through the whole network, these small worlds properties may explain rather intra-component ties since they increase closure within smaller components. On the other hand, during this process, some actors may become more connected than others and their increased degree centrality makes them more visible to the peers' community as a whole and increases their attractiveness and the probability to form a bridging tie with another component. Three hypotheses are tested in order to consider the impact of inventor's network position on their tie formation.

Hypothesis 1a. Inventors will have a higher probability to form a tie, whether bridging or intra-component, with visible inventors having a higher degree centrality level.

In order to test this hypothesis, we compute the "Co-inventors' degree ratio", which is the ratio of the inventor 2's over the inventor 1's degree given that we test the probability for inventor 1 to form a bridging tie.

Hypothesis 1b. Inventors belonging to a larger component than their co-inventors may have a lower probability to form a bridging link.

This hypothesis is tested with the variable "Co-inventors' component size ratio" which is computed as the inventors'1 component size over the inventor's 2 component size.

Hypothesis 1c. Inventors belonging to a large component have a higher probability of forming an intra-component link.

3.2. Geographical proximity

Geographical proximity is at the heart of the network formation issue and often appears as one of the main drivers since many ties take place between inventors located within a short distance (Boschma and Frenken, 2009). But this question is also highly controversial since a number of academics claim that face-to-face and frequent contacts do not need permanent proximity and thus suggest that inventors do not need to be located in the same region. Thus geographical proximity is not a necessary condition for collaboration and networking (Boschma 2005, Torre and Rallet, 2005). This issue is still on the agenda since many studies show the high clustering of networks within regions and large metropolitan areas and no empirical testing enables for the moment to disentangle the respective geographical and density or network effects. Do inventors tend to co-patent more within locations because they are largely concentrated within some locations or do they collaborate because they really need geographical proximity? Some studies have recently shown that proximity increases project efficacy (Fleming, King III and Juda, 2007) and that it is particularly important when the knowledge base is complex (Sorenson, Rivkin and Fleming, 2006). Furthermore, Lobo and Strumsky (2008) show that agglomeration features of metropolitan areas are more important drivers of patent productivity than structural features of inventive networks, suggesting that it is more important for patent productivity to have a high concentration of inventors than having intensively linked actors.

Hypothesis 2a. Inventors may have a higher probability to form a bridging or an intra-component tie if they share a higher geographical proximity.

The variable "Geographic proximity" captures this effect and is equal to 1 if the co-inventors live in the same region.

The impact of geographic proximity may be more important at the beginning of a technological trajectory since the knowledge base is more tacit, less specialized thus needing face-to-face and more frequent contacts (Sorenson, Rivkin and Fleming, 2006). Over time as, the knowledge base becomes more codified and the techniques more standardized, the geographic proximity may become less important (Ter Wal, 2009; Rothaermel and Thursby, 2007).

Hypothesis 2b. Inventors have a lower probability to form a tie when they share similar location as time goes by, that is, as technology becomes less tacit.

This hypothesis is tested using the variable "Period 4 * Geographical proximity" which interacts the last period 2004-2007 dummy and the geographic proximity dummy. We assume this variable to reduce the probability to form a tie within a similar region.

On the other hand, as research processes become more oriented towards exploitation rather than exploration (Nesta and Saviotti, 2007), it may also happen that the knowledge becomes more specific to certain uses and that a higher proximity is needed between knowledge producers and knowledge users especially pharmaceutical or crops companies. This could increase the geographical proximity over time.

3.3. Technological proximity

The connection of previously disconnected components may be explained by various technological strategies. First, they may result from a technological brokerage strategy aiming at connecting previously separated technological communities (Stuart and Podolny, 1999; Burt, 2004) thus leading to cross-disciplinary fertilization (Fleming et al., 2007). Two different stories may explain tie formation. In a more early stage, when basic science needs to be first investigated, inventors may seek specialization effects by becoming connected to partners working in the same technological fields.

Hypothesis 3a. Inventors may have a lower probability to form a tie if they share a dissimilar technological knowledge base.

The variable "Technological distance" is computed considering the IPC classification (3digit) of each patent and applying the Euclidian distance between each couple of inventors' patent portfolio in terms of technological content (i.e. vector 3digit IPC codes).

This effect may even be stronger when inventors share a geographical proximity. In order to test this positive effect, we compute an interaction term between geographical and technological proximity.

Hypothesis 3b. Inventors that are geographically and technologically close have a higher probability over distant inventors to form a tie.

In later stages, especially in exploitation phases, inventors may rather seek complementary knowledge bases. In this case, diversity of the knowledge base will rather become prevalent. For this reason, we expect dissimilarity to have a positive impact on forming ties as times go by.

Hypothesis 3c. Inventors may have a higher probability to form a bridging tie if they share dissimilar technological knowledge base at later stages of the network.

This hypothesis is tested using the variable "Period 4 * Technological distance" which interacts the last period 2004-2007 dummy and the variable "Technological distance". We assume this variable to reduce the probability to form a tie within a similar region.

3.4 Strategic behavior of network applicants

If the network formation and structure is explained by the inventors' characteristics, the role of patent applicants is most often ignored although co-patenting may reveal some applicants' strategies in terms of knowledge production, diffusion and exploitation. Fleming et al. (2007), for example, highlight IBM Almaden Valley Labs' structural role in giant component formation as IBM highly invested in research and offered a doctoral program for Stanford University students thus favoring the connection between IBM and their doctoral students future appointments.

These bridging and intra-component links may reveal the applicants' strategies in terms of cumulative and closure effects. If an applicant is active in both components that become connected through co-patenting, this reveals a cumulative effect, that is, the applicant's closure. On the other hand, if the bridging tie connects components in which the applicant was only active in only one of them or even in none of them, this reveals a strategic intention of bridging separate communities.

Hypothesis 4a. Inventors may have a higher probability to form a bridging tie if the applicants have a cumulative strategy and try to connect components in which it is already active.

Hypothesis 4b. Inventors may have a higher probability to form an intra-component link if the applicant has a closure strategy

Hypothesis 4c. Inventors may have a higher probability to form a bridging tie if the applicants have a selective strategy and tries to connect become connected to a component in which it is not already active

(This part will be developed in the full version of the paper)

In order to test these hypotheses, we consider if applicants are already active in one or both components that become connected. This will be checked, both globally considering the overall network and locally around the focal network, considering her ego-network (defined considering a geodesic distance equal to three). This latter option should be more appropriate since first it permits to control for the size of component, second it assumes an imperfect information distribution within a component.

4. Method and estimation procedures

4.2. Dependent variable, controls and estimation

In order to study the dyad formation, we follow other studies that estimate the likelihood of a tie formation (Sorenson, Rivkin and Fleming, 2006). We first computed all existing and potential dyads between two pairs of inventors. The existing dyads are divided in four categories detailed in Table 1. All the possible and realized dyads generate around four millions observations and the realized links represent only a marginal portion of all possible links. Since this gap raises important difficulties of estimation, we adopted a case-control approach.

The case-control approach, to analyze the formation of co-inventor links, is based on the following procedure. We begin by considering any inventor forming a new bridging link (a similar approach is applied for the intra-component links) as well as her co-inventor. In this case, the dependent variable, a bridging link takes a value of "1" for these cases to denote a realized link. In addition, we consider for each of both "real" co-inventors, five possible but not realized co-inventors which provide five controls for each co-inventors and in sum ten potential co-inventors for each realized dyad. These bridging links are set to zero for these control cases. A similar procedure is applied for intra-component links. We test the probability of forming a tie using a conditional logistic regression estimation procedure.

The first model estimates the probability for an inventor in forming a bridging link and the second model estimates the probability of forming an intra-component link as a function of the inventors' network effects, geographical and technological proximity and applicants' strategy. The odds-ratios are given in the table 2 and 3 in order to ease the analysis. Besides the explanatory variables, a number of controls are introduced. Following Powell et al. (2005), we introduce a linear trend variable, named *Timeline* as well as a *Period* variable in order to capture some changes occurring over time. The *Timeline* variable is computed as the current year of patent application and 1992. The whole period has been divided in four phases and a dummy has been introduced in order to capture the last phase from 2003 and 2007. Since most inventors are located in the Paris region Ile-de-France, Rhône-Alpes and Alsace, a dummy has been introduced in order to control for specific effects of these locations.

4.1. Estimation results

The estimations provide support for the importance of the inventors' relative position in terms of degree centrality in explaining the tie formation. The co-inventors degree ratio is positive for the bridging as well as intra-component links suggesting that inventors tend to create links with actors that are more densely connected either in another component or within their own. Contrary to our hypothesis 1b, they are not attracted to larger components when they form bridging ties, suggesting that they tend to be connected to smaller ones, probably because most of the bridging links consist of a smaller component being connected to a larger one thus forming a giant component in the end. The intra-component links do not necessarily occur within the largest components, contrary to our expectations. It is significant only for model 1 and slightly for model 2 when only social proximity is considered. It does not remain significant once geographical and technological variables are introduced suggesting that the size of the component is not determining in network formation.

Model 3 tests both geographical and technological proximity as well as their interaction term. Both factors are significant suggesting that links have a higher probability to be formed among inventors that share the same location and that have similar technological specialization. The interaction term is only slightly significant for intra-component links suggesting that both effects are increased when geographical and technological proximity occur at the same time.

Model 4 and 5 introduce the effect of the fourth period and its interaction term with the geographical and technological proximity. The results are not totally evident to explain. The interaction term has a negative and significant effect, reinforcing the technological proximity. They nevertheless may suggest that technological proximity still plays a role even in latter periods for the bridging ties. This could indicate that bridging ties do not aim at opening the knowledge base by getting connected to components with different knowledge bases. It rather suggests that bridging strategies aim at increasing technological specialization. The geographical proximity does not play or role, it does not seem to decrease over time. For the intra-component links, the picture is completely the opposite. The interaction term is very positive and strong indicating that contrary to our expectations; the geographical proximity is still a dominant factor in intra-component links. In order to understand if genomic field has reached a certain maturity stage or not and control the evolution of tie formation, we should control in the next version of the paper for technological diversity and maturity in the network.

5. Conclusion

This paper investigates the dynamics of network formation in genomics patenting in France over the last two decades and it tests the respective role of cumulative mechanisms and actor strategies in determining the evolution of inventor networks. To do that, we have considered two different types of links (a link bridging two components and an intra-component link) and started to explore the determinants of their formation. First results show common elements and some differences between their determinants. Concerning the former, we found: first, inventors tend to create links with actors that are more densely connected either in another component or within their own; second geographical distance seems to play a mayor role as well as technological distance. Concerning the role played by this latter, it emerges an important difference between the formation of bridging links relative to intra-component one. Being near in the technological space is always important for intra-component links, while technological distance plays a role in bridging link formation only in most recent years of the period under analysis.

References

Amburgey, T, Al-Lahan A., Tzabbar D., Aharonson B. (2008), The structural evolution of multiplex organizational networks: research and commerce in biotechnology", Advances in Strategic Management, *Network Strategy*, Vol 25, 171-209.

Boschma R., Frenken K. (2010), The Spatial Evolution of Innovation Networks: A Proximity Perspective in Boschma, R. and Martin R.L. (Eds), Handbook of Evolutionary Economic Geography, Edward Elgar, Forthcoming

Baum, J. A. C.; Shipilov, A. V. & Rowley, T. J. Where do small worlds come from? *Ind Corp Change*, 2003, 12, 697-725

Boschma R., (2005), Editorial: Role of Proximity in Interaction and Performance: Conceptual and Empirical Challenges, *Regional Studies*, vol. 39(1), pages 41-45

Breschi, S & Lissoni, 2001. Knowledge Spillovers and Local Innovation Systems: A Critical Survey, *Industrial and Corporate Change*, Oxford University Press, vol. 10(4), pages 975-1005

Burt, R. (2004), "Structural holes and good ideas." *American Journal of Sociology*, 110: 349-399.

Fleming, L., Frenken, K. (2007), The evolution of inventor networks in the Silicon Valley and Boston regions. *Advances in Complex Systems* 10 (1): 53-71.

Glucker, J. (2007) Economic geography and the evolution of networks. *Journal of. Economic Geography* 7, 619-634

Jaffe, A. B.; Trajtenberg, M. & Henderson, R. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations The Quarterly Journal of Economics, The MIT Press, 1993, 108, 577-598

Powell, W. W.; White, D. R.; Koput, K. W. & Owen-Smith, J. Network (2005), Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences The *American Journal of Sociology*, The University of Chicago Press, 110, 1132-1205

Rallet A. and Torre A. (2005), "Proximity and Localisation",Regional Studies, 39:1, pp. 37-59, 2005

Sorenson, O.; Rivkin, J. W. & Fleming, L. (2006), Complexity, networks and knowledge flow *Research Policy,* 35, 994 – 1017

Stuart, T. and J. Podolny, (1999), "Positional Causes and Correlates of Strategic Alliances in the Semiconductor Industry." Pg. 161-182 in *Research in the Sociology of Organizations*, edited by Steven Andrews and David Knoke. Greenwich, CT: JAI Press

Ter Wal A. L.J. (2009), The Spatial Dynamics of the Inventor Network in German Biotechnology: Geographical Proximity Versus Triadic Closure, Economic Geography Research Group Working Paper Series No. 02.09

Watts and Strogatz, 1998 D.J. Watts and S.H. Strogatz(1998), Collective dynamics of 'small-world' networks, *Nature* 393, pp. 440–442.

Table 2. Conditional Logistic Regression for Bridging links – Odds Ratio

	Model 1	Model 2	Model 3	Model 4
Co-inventors' degree ratio	1.095***	1.089***	1.090***	1.088***
	(0.021)	(0.021)	(0.024)	(0.024)
Co-inventors' component size ratio	1.007*	1.007*	1.007*	1.007+
	(0.003)	(0.003)	(0.003)	(0.003)
Co-inventors' Technological distance			0.315***	0.636
			(0.134	(0.301)
Co-inventors' Geographical proximity			9.807***	8.767***
			(3.149)	(3.118)
Technological distance * Geographical proximity			0.6741	0.652
			(0.372)	(0.362)
Period 4				1.376
				(0.929)
Period 4 * Geographical proximity				1.337
				(0.464)
Period 4 * Technological distance				0.121***
				(0.08)
Timeline (year – 1992)	1.032	1.013	1.016	1.059
	(0.058)	(0.058)	(0.062)	(0.092)
Paris Region		0.3362***	0.677	0.655
		(0.083)	(.185)	(0.179)
Rhone – Alpes Region		0.530	1.656	1.625
		(0.167)	(0.571)	(0.562)
Alsace Region		0.155***	0.359+	0.325+
		(0,084)	0,205	(0,191)
# observation	2111	2111	2111	2111
LR chi2	26.17	65.87	254.08	265.07
Pvalue	0.0000	0.0000	0.0000	0.0000
Pseudo R2	0.0222	0.0559	0.2157	0.2251

Table 3. Conditional Logistic Regression for Intra-component links – Odds Ratio

	Model 1	Model 2	Model 3	Model 4	Model 5
Co-inventors' degree ratio	1,160***	1,137**	1,151**	1,151**	1,164***
	(0,048)	(0,046)	(0,056)	(0,056)	(0,058)
Co-inventors' component size	0,996*	0,997+	0,998	0,998	0,997
	(0,002)	(0,002)	(0,002)	(0,002)	(0,002)
Co-inventors' Technological distance			0,042***	0,053***	0,066***
			(0,025)	(0,032)	(0,039)
Co-inventors' Geographical proximity			5,465***	4,230***	5,659***
			(1,853)	(1,516)	(1,978)
Technological distance * Geographical proximity			3,458+	3,531+	3,107
			(2,514)	(2,584)	(2,250)
Period 4				0,449	0,417
				(0,369)	(0,343)
Period 4 * Geographical proximity				2,627+	2,828*
				(1,310)	(1,406)
Period 4 * Technological distance				0,479	0,419
				(0,414)	(0,363)
Timeline (year – 1992)	1,062	1,038	1,077	1,123	1,146
	(0,065)	(0,065)	(0,076)	(0,116)	(0,118)
Paris Region		0,168***	0,306**	0,328**	
		(0,059)	(0,125)	(0,135)	
Rhone – Alpes Region		0,123***	0,366	0,353	
		(0,089)	(0,276)	(0,265)	
Alsace Region		(0,139)	0,419	0,425	
		0,101**	(0,313)	(0,318)	
# observ.	1664	1664	1662	1662	1662
LR chi2	19.27	81.37	281.13	286.16	275.47
Prob > chi2	0.0002	0.0000	0.0000	0.0000	0.0000
Pseudo R2	0.0199	0.0841	0.2917	0.2969	0.2858

+ 0.10 * 0.05 ** 0.01 *** 0.001