



HAL
open science

Network Formation and Geography: Modelling Approaches, Underlying Conceptions, Recent and Promising Extensions

Corinne Autant-Bernard, Cilem Selin Hazir

► **To cite this version:**

Corinne Autant-Bernard, Cilem Selin Hazir. Network Formation and Geography: Modelling Approaches, Underlying Conceptions, Recent and Promising Extensions. 2013. halshs-00803536

HAL Id: halshs-00803536

<https://shs.hal.science/halshs-00803536v1>

Preprint submitted on 22 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

WP 1312

**Network Formation and Geography: Modelling Approaches,
Underlying Conceptions, Recent and Promising Extensions**

Corinne Autant-Bernard, Cilem Selin Hazir

March 2013

GATE Groupe d'Analyse et de Théorie Économique Lyon-St Étienne

93, chemin des Mouilles 69130 Ecully – France

Tel. +33 (0)4 72 86 60 60

Fax +33 (0)4 72 86 60 90

6, rue Basse des Rives 42023 Saint-Etienne cedex 02 – France

Tel. +33 (0)4 77 42 19 60

Fax. +33 (0)4 77 42 19 50

Messagerie électronique / Email : gate@gate.cnrs.fr

Téléchargement / Download : <http://www.gate.cnrs.fr> – Publications / Working Papers

Network Formation and Geography: Modelling Approaches, Underlying Conceptions, Recent and Promising Extensions

Corinne AUTANT-BERNARD and Çilem Selin HAZIR

Université de Lyon, Lyon, F-69007, France ; Université Jean Monnet, Saint-Etienne, F-42000, France ; CNRS, GATE Lyon Saint-Etienne, Ecully, F-69130, France.

corinne.autant@univ-st-etienne.fr, cilem.selin.hazir@univ-st-etienne.fr

Abstract Due to the strong polarisation of economic activities in space and rise in collaborative behaviour, increasing attention has recently been devoted to the relationship between geography and network formation. The studies conducted on this topic reveal a high variation in terms of methodologies. Putting special emphasis on R&D networks, the aim of this chapter is to review the different methods and assess their ability to address the issues raised by the relationship between network and space. We first discuss the different facets of the relationship between geography and networks. Then, we detail the methodological approaches and their capability to test each effect of geography on network formation. We argue that the effect of distance on dyads have received the major attention so far, but the development of block modelling and top-down approaches opens new research perspectives on how distance or location might affect formation of more complex structures. Moreover, recent improvement in temporal models also offers opportunities to better separate spatial effects from that of influence over time.

Keywords : Social Network Analysis, Stochastic Actor-based Models, Gravity models, Complex Network Analysis, ERGM, Preferential Attachment Model

JEL Codes C18, O33, R11

1. Introduction

In the field of economics, the relationship between geography and network formation attracts attention in order to understand how knowledge flows in a space of social interactions relate to regional growth and innovation. So far a number of studies have been conducted to elucidate this relationship. Even a glimpse on these studies reveals a high variation in terms of methodologies.

On the one hand, this variety stems from the fact that the term “*geography*” contains a number of meanings in it. Sometimes geography is associated with physical separation, sometimes it refers to locations as a material and relational context for economic action, and sometimes geographical units themselves are considered nodes in a network. The way it is conceived, in turn affects the way it is related to network formation and constrains model choices as some models are not capable of testing all kinds of effects. On the other hand, the variety in methodologies results from addressing the same phenomenon; i.e. formation of a network through different analytical perspectives.

From a “learning perspective” these differences enclose invaluable information on the evolution of the way that the research community has conceived and addressed the geographical dimension of network formation, and on possible future directions. In this regard, this chapter will try to disclose this information by elaborating how different meanings associated to geography can yield different conceptualizations of geography-network relationship. Hence in Section 2 we will address alternative ways of relating geography to network formation. In Section 3, we will try to identify main distinctions between different methodologies and compare models that are widely used in the study of spatial dimension of R&D networks. Our aim here is not to provide a full-fledged list and a hierarchy of network formation models but rather to highlight main differences in analytical approaches putting emphasize on their ability to address the issues raised by the relationship between network and space. Finally, we will review some recent methodological advancement that looms large regarding their potential future contributions to understand knowledge flows in space.

2. Relating Geography to Network Formation

2.1 A tie covariate: physical distance

One of the meanings associated to geography is the physical distance, which is the relative position or physical separation of two entities. Under this definition, space is perceived to be homogenous and exogenous to the network formation process due to the fact that regardless of the configuration of the network, the physical distance among nodes remains unchanged. Then, the role of geography is conceptualized as the effect of an attribute of a possible tie; i.e. the length of a tie.

High levels of this attribute is hypothesized to have a negative effect on the utility out of being connected¹ due to the fact that there exists a *tacit* component of knowledge (Polanyi 1966) and some interaction is necessary for its transmission.

¹ As shall be seen in the succeeding section, this utility either refers to a utility obtained out of a tie (see binary choice models), or to the utility out of the overall network (see ERGM).

Therein, physical proximity is considered to be a facilitator of face-to-face interactions, which in turn eases the transmission of tacit knowledge (Feldman and Florida 1994) and hence increases the utility of being connected². Also, physical proximity is assumed to increase this utility via enabling cross-fertilization of ideas (Feldman and Florida 1994) and timely inflows of information (Feldman 1993) and by decreasing the cost of collaboration (Hoekman et al. 2009).

However, the fact that physical distance is just one of the many dimensions of separation (Boschma 2005) and in particular the embeddedness of economic relations between firms and individuals in social relations (Granovetter 1985) has modified this hypothesis. Thus, it has become a matter of interest to know whether physical distance still plays a role on the utility of being connected when the effects of other dimensions of separation are controlled for.

2.2 A node covariate: local context

Another meaning that is associated to geography is the physical context that economic agents are embedded in. Once the context that embraces networking agents is taken into account, then the network becomes embedded in a physical space. One way to relate this embeddedness to network formation is to consider the physical space as an exogenous setting, which affects the attractiveness of the organizations as potential partners or their capacity to establish connections. In that case, the role of geography is conceptualized as the effect of a node attribute on network formation. In the literature, this effect is formulated in a number of ways such as the effect of agglomeration economies, knowledge externalities, system of innovations, or innovative “milieu”.

Although, considering geography as an exogenous node attribute simplifies the analytical processes to study network formation; obviously the local processes and network processes are not mutually exclusive. On the one hand, the black-box of advantages that a location provides might also include the outcomes or impacts of network activity of its constituents. On the other hand, some local processes might not only work through increasing node attractiveness or capacity but also through creating tie dependence as will be discussed in the sequel.

² However, if proximity is often associated with the tacit dimension of knowledge, we must avoid an overly simplistic view (Massard and Mehier, 2009). There are probably complementarities between tacit and codified knowledge, any two being transmitted both locally and remotely. The link between proximity and knowledge can then lie in the way of combining the tacit and codified nature of knowledge.

2.3 A factor affecting tie dependence: physical distance and local context

Pattison and Robins (2002) argue that each network tie could be associated with a “social locale”, which refers to “a complex relational entity that links the geographical, social, cultural and psychological aspects of the context for social action”. They argue further that these social locales overlap with each other due to the fact that “the outcome of processes in one locale may have some impact on processes within another locale”. Therein, a local context might be considered as a joint social locale for ties created within it, as they all share a number of intermingled local processes such as social, economic, political, historical processes. The outcomes of these processes might be heterogeneous across space and they may create, enhance or even dampen dependencies among ties. Similarly, being spatially proximate could be associated with overlaps in social locales as being spatially proximate might mean sharing similar local features.

Hence, in this case the role of geography can be conceptualized as the effect of tie dependence on network formation. Unlike considering the role of geography as the effect of a tie attribute, in this conceptualization the specific role played by distance is not disentangled from the role of other types of proximities or processes that co-exist or interact with geographical proximity.

2.4 Regions as nodes themselves

As a matter of fact, geographical units may themselves constitute the nodes in a network. In the case of networks representing economic, regions symbolize the aggregate behaviour of individuals. Hence, all three types of roles discussed above might be relevant to study the inter-regional networks. The role played by the distance between two regions or existence of a common border might again be considered as an exogenous tie property. Regional properties that might affect the aggregate performance of individuals can be considered as exogenous node attributes under the assumption that network processes and these properties are mutually exclusive. Finally, contiguity or co-location in a wide geographic area can be conceptualized as a factor affecting overlaps in social locales.

3. Approaches to model network formation

Networks attract attention from a wide range of fields like medicine, biology, computer science, sociology, political science, economics, etc. Accordingly, a number of different analytical approaches have been suggested to model their formation. A major distinction among these approaches stems from considering

the network as an outcome of “choice” or “chance”. In the first view, formation of a network is explained on the basis of individual incentives (costs and benefits) (Jackson and Wolinsky 1996). A number of strategic and game theoretic models have been developed along this view. On the other hand, graph-theory has bestowed various random graph models in line with the second view, where the observed network is considered as just one realization among all possible network configurations. Beside random graphs, complex network analysis has been developed along the same line. Finally, the usual econometric models and spatial econometrics have also been applied to study the network, where both views are in play.

These approaches may also be classified into two as static approaches and dynamic approaches. The former works on a snapshot of the network; whereas the latter considers the evolution of the network in time. Among those, some models allow creation of new nodes in time as in the case of preferential attachment model (Barabási and Albert 1999). Some others allow studying the dynamics stemming from creation and dissolution of ties among a fixed set of nodes in time as in the case of stochastic actor-based models (Snijders et al. 2010).

As a third classification, these approaches can be considered in two groups as top-down approaches and bottom-up approaches. Top down approaches focus on the topology of the network as a whole and try to identify global features rather than modelling the network on the basis of individuals. Complex network analysis or block modelling (Nowicki and Snijders 2001), where the aim is to identify groups, members of which are equivalent in terms of their connection patterns may illustrate this approach. On the other hand bottom-up approaches focus on processes taking place in components of the network. Therein, a further distinction can be made among bottom-up approaches with respect to the types of components that they focus. In some approaches the network configuration is explained by focusing on the behaviour of actors, ex: stochastic actor-based models (Snijders et al. 2010). Whereas in some others the focus is either on formation of a single tie or a local pattern (a subset of ties).

Another distinction among these approaches could be made with respect to underlying assumptions on tie dependence. Some models base on the assumption that the stochastic processes behind formation of ties work independently. Some others assume that the outcomes of these stochastic processes are correlated. Finally a third group assumes that some ties are realized jointly through the same stochastic process.

In the sequel, we will focus mainly on the empirical studies that investigate the role of geography in R&D networks. We will discuss them under three headings: network as the equilibrium of choices; network as an outcome of choice and random effects, and network as an outcome of a random process. We will try to highlight the differences in the analytical process among these models in terms of the above-mentioned criteria and their capacity to handle alternative ways of relating geography to network formation

3.1 Network as the equilibrium of choices

As mentioned earlier the game theoretic approach considers the network as the outcome of individual choices. Among these models the seminal work by Jackson and Wolinsky (1996) has considerable influence on both theoretical and empirical work on the geographical dimension of R&D networks. Their model, known as the connections model, explains the formation of a network on the basis of individual incentives (costs and benefits) and bases on the idea that agents do not only benefit from those they are linked directly; but also from those they are linked indirectly. The benefit they can obtain from others decreases with distance; but direct links are costly implying a trade-off between the benefits and costs of a direct link.

The spatial extensions of this model is provided by Johnson and Gilles (2000) and Carayol and Roux (2007). In these extensions the role of geography is investigated in a static network, where the number of nodes is fixed. Geography is considered as the geographical distance and its role is hypothesized as an exogenous factor affecting the cost of maintaining a link. Based on this conception on the geography-network formation relationship, these theoretical models suggest that for a wide range of intermediary values of decay in transmission of knowledge, a particular stable network structure called “small world” emerges. Carayol and Roux (2007) also provide some empirical evidence by fitting the model to actual co-inventions that took place during 1977-2003 with at least one inventor located in France.

3.2 Network as an outcome of choice and random effects

While in the game theoretic models the network is considered as the equilibrium of individual choices, in some statistical models used to study connections among nodes we see an expression of the utility that an individual can obtain out of its choice and some notion of randomness in making that choice. In the sequel, these models will be explained briefly and their capacity to integrate the geographical dimension will be discussed.

Binary Choice Models The use of Binary Choice Models illustrates the application of usual econometric tools to study network formation (Geuna 1998, Powell et al. 2005, Mairesse and Turner 2005, Autant-Bernard et al. 2007, Paier and Scherngell 2008). These models aim at explaining the factors that affect realization of a single tie; hence they analyse formation of a network by focusing on its smallest unit. Factors that are symmetric for a pair of nodes, i.e. tie attributes, are the easiest ones to test with these models. Some practical problems arise in studying the effect of node attributes since the explanatory variables have to be symmetric and hence insensitive to the changes in the order of indexation. Finally, these models allow studying the effect of the observed network configuration on

tie formation but under the assumption that it is an exogenous factor. This stems from the fact that in these models realization of a tie is supposed to be a Bernoulli process, meaning that ties are realized independently of each other.

Therein, the capacity of Binary Choice Models to investigate the role of geography mainly lies in the ability to study how physical distance affects the probability that a tie is created given the effect of other factors. This ability complies with the research interest to demarcate the role of geographical proximity from that of other proximity dimensions. In these models geography may also be included as a node attribute as long as they are defined symmetrically for the pair of nodes. Finally, due to the tie independence assumption, with these models it is not possible to study the role of geography in terms of tie dependence.

Poisson Regression Models and Gravity Models The analytical process and the assumptions in Poisson Regression Models are the same as those in Binary Choice Models except for the fact that the objective is to explain the intensity of interaction among a pair of nodes rather than its existence (Powell et al. 1996). Hence, they allow studying the role of geography on the intensity of interactions, where this role could be introduced as a tie or node property (Mairesse and Turner 2005, Frachisse 2010). Once distance is accounted for, Poisson models can be interpreted as gravity models. As in the case of Poisson Regression Models, the objective in Gravity Models is to explain the strength of interaction among two spatial units. Hence, the approach undertaken to explain for the network builds upon ties among pairs. This type of models can be applied to individual choices or aggregated behaviour. It is worth noticing however that much attention has been devoted so far to study inter-regional networks, hence focusing on aggregated data.

The use of these models illustrates an application of spatial analysis techniques to study network formation. The earlier studies using Gravity Models assume that the stochastic process behind tie formation works identically and independently; i.e. any pair of ties, among the same pair of nodes or not, are independent (Ponds 2006, Maggioni 2007, Scherngell and Barber 2009, Hoekman et al. 2010). More recent applications (Scherngell and Lata 2011) take the spatial autocorrelation among flow residuals into account and corrects for this by using eigenvector filtering. Hence, the extension with spatial filtering rests upon weaker assumptions on tie dependence since it handles the correlation among ties sharing the same node.

As Gravity Models include two mass terms and a separation function; they allow studying the role of geography as a node itself with some attributes and as a tie attribute. The extensions dealing with spatial autocorrelation might allow controlling for correlations among intensity of interactions resulting from the topology of regions. Hence, the specific role played by the physical distance might be identified better as suggested in Chapter 11 of this book.

Stochastic Actor-based Models Stochastic Actor-based Models are statistical models to study tie dynamics in networks of fixed size (Snijders et al. 2010). As the name implies they focus on the behaviour of actors and model the formation of the network by means of changes that actors make in their outgoing ties. These changes are explained by means of two functions. The former is the rate function

showing the frequency at which a change occurs. Whereas, the latter refers to the objective function, which shows the probabilities of alternative courses of action given the opportunity to make a change. This function is expressed in terms of “effects”, which are tendencies (like reciprocity, closure, multi-connectivity etc.) taking place locally³. Both functions may depend on network position of actor and some actor attributes.

These models assume that actors act independently; hence the changes they make are not coordinated yet sequential. However, as the outcomes of their decisions change each other’s environment, in time their actions depend on each other. Thus, unlike Binary Choice Models, where for each pair of agents the rest of the network is considered exogenous simultaneously; the sequential nature of Stochastic Actor-based Models allow handling dynamism in choices and dependencies on the environment.

Geographical dimension might be introduced in these models through both the rate and the objective function. A rate function differentiated with respect to location of actors might enable spatial heterogeneity in frequency of tie changes. On the other hand, the objective function might be modified either by integrating the distance as a dyadic covariate (Ter Wal 2013), or location as a node attribute (Balland 2012) which in turn might be used to study the effect of co-location and some network effects arising from being co-located.

3.3 Network as an outcome of a random process

As mentioned earlier the graph theoretic approaches consider the observed network as an outcome of a random process. Hence, these approaches do not base on utility functions of micro agents but on the distribution of probabilities. Nevertheless, it should also be noted that although a utility function is not specified in these models, the distribution of probabilities can be constrained using a theoretical basis on preferences of agents⁴. Below, graph theoretic approaches used to study geographical dimension of R&D networks are discussed.

Exponential Random Graph Models (ERGM or p*) ERGMs are (Frank and Strauss 1986; Wasserman and Pattison 1996) more recent types of random graph

³ These effects are similar to the “local configurations” in Exponential Random Graph Models that will be discussed in the sequel.

⁴ As shown by Park and Newman (2004) random graph models can be expressed as a constrained maximum entropy problem; which maximizes the entropy in the probability distribution of observing a particular network configuration. In the earlier random graph models (Erdős and Renyi 1959) the problem is constrained only by the number of the edges in the network and a probability distribution which assigns the same probability to all networks with the same number of edges is obtained. However, in more recent models as shall be seen in subsection on Exponential Random Graph Models, the preferences of actors for homophily, central agents, closure, etc. can be used as additional constraints by defining local configurations accordingly.

models. They allow studying networks with a fixed set of nodes. However, as will be discussed in Section 4, temporal extensions that allow dynamism in terms of tie creation and dissolution have recently become available. The idea behind these models is that the observed network is just one realization of all possible configurations of connections among a given set of nodes. Hence, as stated by Cranmer and Desmarais (2010), there is a “conceptual leap” from the Binary Choice and similar models to ERGM. While in the former the vector of interest is a series of values drawn from a univariate distribution; in ERGM it is considered as a single draw from a multivariate distribution. This leap allows relaxation of the tie-independence assumption and provides ERGMs capacity to tackle with even complex dependence structures among ties (see realization dependence assumptions by Pattison and Robbins (2002)).

An ERGM explains the formation of a network by means of local configurations, which are some small and regular patterns. Among all possible networks configurations it gives a higher probability to those that are similar to the observed network in terms of these small structures. In defining these local configurations, ERGM is capable of differentiating ties and nodes with attributes (Robins et al. 2007).

Therein, the capacity of ERGMs to investigate the role of geography is three folds. First, it may be studied as the role of physical distance by means of a distance interaction function (Daraganova et al. 2012). Second, geography may be included as a node attribute. Third, geography can be considered as a spatial setting that imposes limits on tie dependence, hence on local configurations (Pattison and Robbins 2002). The studies by Broekel and Hartog in Chapter 4 and Hazir in Chapter 13 illustrate the applications of these models on R&D networks.

Preferential Attachment Model Preferential Attachment Model (Barabási and Albert 1999) is a graph theoretical model explaining dynamic networks with growing number of nodes. The model in its original form explains the formation of a network as a process where the degrees of existing nodes increase proportional to their magnitude and result in a scale-free degree distribution. Hence, it considers a single factor; i.e. degree affinity of agents, to explain for the network via explaining one of its macro properties; i.e. its degree distribution. The extension by Vinciguerra et al. (2010) integrates the effect of geographical distance and co-location in the same country to the probability that a node receives connections as the network grows.

Complex Network Analysis (CNA) While the models reviewed so far aim at explaining the formation of a network by means of micro processes, Complex Network Analysis focuses the overall topological structure of complex networks. Hence CNA aims at identifying and explaining key global features like degree distribution, diameter, clustering, and communities.

A number of studies revealed that R&D networks display a scale-free degree distribution, “small-world” property in terms of diameter and high “clustering” (Goyal et al. 2006, Newman 2001, Gay and Dousset 2005). On the one hand, theoretical models (Johnson and Gilles 2000, Carayol and Roux 2007), the above men-

tioned spatially extended preferential attachment model, and possibly ERGMs illustrate how a model explaining the effect of geography on micro processes can also explain for these global properties. On the other hand, CNA adopts a macro perspective to study the spatial dimension of such properties. The study by Barber and Scherngell in Chapter 10 illustrates studying the heterogeneity in the spatial configuration of communities in an R&D network. Whereas, the study by De Montis et al. in Chapter 3 illustrates the use of CNA to investigate whether similar geographical contexts give rise to similar global network properties or not.

4. Conclusions and Future Directions

In this chapter, we considered the relationship between geography and network formation but our focus was on how to investigate this relationship. Hence, we reviewed different meanings of geography and different conceptualizations of this relationship. Then we provided an overview on different approaches through which network formation is explained. Our aim was neither to provide a complete list or a hierarchy of network formation models nor to identify best models. Rather we were interested in two aspects. First, leaving all the practical issues and formal definitions of models aside, we aimed at identifying the grand avenues that a researcher can follow in studying network formation. We identified that whether to consider it as an outcome of choice or chance; whether to consider it as a dynamic or a static process; whether to explain it from bottom-up or top-down; whether to study its complex interdependencies or simplify it are the major decisions to be made by the researcher in making a model choice. Second, all these choices suggest a different capacity to study the role of geography. Hence, we reviewed applied studies with a particular interest on those on R&D networks to highlight these analytical differences, the evolution of analytical frameworks (if any) and to identify future directions.

One of the main conclusions that could be derived from this review is that so far the research community made use of mainly bottom-up approaches to study the role of geography in formation of R&D networks. In other words, the emphasis is given to explain how geography affects the formation processes at the micro level. Although global topological features of these networks have attracted attention, spatial heterogeneities in these global features or heterogeneities in spatial patterns of components of networks have received less attention. Apart from those techniques used by Barber and Scherngell in Chapter 10 and by De Montis et al. in Chapter 3; block modelling might also be used to study the relationship between geography and network components, members of which are equivalent in terms of their connection patterns.

Another conclusion could be derived on the evolution in the analytical processes that are adopted to study the role of geography in formation of R&D networks. While there is not a clear cut distinction, it is noteworthy that the community has recently shown interest in models that can allow dependence among ties.

This enables demarcating the effect of dependence from other factors of interest; hence improves estimates for the role of geography. Furthermore, the ability of these models to handle tie dependence result in models which can explain the global topology of the network (such as clustering, degree distribution, etc.) as well as the local process in focus. However, the relationship between geography and tie dependence is far from being exploited. So far the effect of distance on dyads have received the major attention, leaving how distance or location might affect formation of more complex structures than dyads left aside.

A third conclusion stems from the temporal dimension of networks. As a matter of fact most applied studies consider an R&D network as a static object, where neither new nodes are added nor ties created or dissolved. Applications of stochastic actor-based models relaxed this assumption and considered the tie dynamics among a fixed set of nodes. These models indeed possess a capacity to analyse not only the determinants of tie formation but also tie dissolution by means of an endowment function (Snijders et al. 2010). Hence, these models may well be used to study the role of geography on tie dissolution. Although not applied to study the geographical dimension of R&D networks some recent temporal extensions of ERGM also suggest similar possibilities. Among these Hanneke et al. (2010) provides Temporal Exponential Random Graph Model (TERGM), which allows studying the evolution of a network of fixed size. Whereas, Krivitsky and Handcock (2010) enables separating tie formation and dissolution processes in a TERGM.

Apart from these some other model extensions suggests additional explanatory capacity for the field. Among these, extension of ERGMs for valued networks (Krivitsky 2012) stands as another tool to study the effect of geography on the intensity of connections, which has been studied so far by means of Poisson regression models and gravity models. The ability of this tool to handle tie dependence might be useful for better treatment of network effects and demarcate the role of geography more properly. In addition to that, Steglich et al. (2010) extended Stochastic Actor-based Models to distinguish partner selection from social influence in a dynamic network. This extension basis on the idea that two actors showing the same behaviour might be collaborating due to similarity in their behaviour, or one gets similar to the other as a result of being connected. The ability to separate those two processes might be valuable in better demarcation of spatial effects from that of influence over time.

In addition, by improving our understanding of network formation and evolution, all these developing techniques may also contribute to a better comprehension of the mechanisms that generate network outcomes. A growing literature tries to understand how some of the particular topological network properties (such as density, clustering, connectivity of the network, degree distribution of nodes or degree assortativity) influence economic performances at the regional level (Breschi and Lisson, 2011, Crespo, Suire, Vicente, 2013). However, as argued by Ahuja, Soda and Zaheer (2012), “without a comprehension of the logic that drives network creation, scholarly understanding of their outcomes remains incomplete” (p. 34). In particular, as it is difficult to identify whether the network structure im-

plies the outcome or the reverse, we have to consider both aspects together. To this respect, the contributions of spatial econometrics to the field of network analysis may extend beyond gravity models as suggested by (Autant-Bernard 2012). Spatial tools can indeed provide valid instruments allowing endogenous effects to be separated from exogenous ones (see for instance Bramoullé et al. 2009). In the same line, the temporal extensions of the above reviewed network approaches are also very promising in order to cope with this causality problem.

Finally, it is a matter of fact that model choices are strongly constrained by the nature of data and data availability. Assumptions of a model might be severe or reasonable depending on the nature of the data and on the properties of the economic process through which it is generated. Hence, there is no one-for-all answer on how to study the spatial dimension of network formation.

Acknowledgements

This research has received funding from the European Community's Seventh Framework Programme under grant agreement n° 266834.

References

- Ahuja G, Soda G, Zaheer A (2012) The genesis and dynamics of organizational networks. *Organization Science* 23: 434-448 <http://didattica.unibocconi.it/mypage/upload/>
- Autant-Bernard C et al (2007) Social distance versus spatial distance in R&D cooperation: Empirical evidence from European collaboration choices in micro and nanotechnologies. *Papers in Regional Science* 86(3): 495-519
- Autant-Bernard C (2012) Spatial Econometrics of Innovation: Recent Contributions and Research Perspectives. *Spatial Economic Analysis* 7(4): 405-421
- Balland PA (2012) Proximity and the Evolution of Collaboration Networks: Evidence from Research and Development Projects within the Global Navigation Satellite System (GNSS) Industry. *Regional Studies* 46(6)
- Barabási AL, Albert R (1999) Emergence of Scaling in Random Networks. *Science* 286(5439): 509-512
- Boschma RA (2005) Proximity and Innovation: A critical Assessment. *Regional Studies* 39(1), 61-74.
- Bramoullé Y, Fortin B (2009) The Econometrics of Social Networks, Centre Interuniversitaire sur la Risque, les Politiques Économiques et l'Emploi Working Papers 09-13. Available at: http://www.cirpee.org/fileadmin/documents/Cahiers_2009/CIRPEE09-13.pdf
- Breschi S, Lenzi C (2011), Net and the city. Co-invention networks and the inventive productivity of US cities, mimeo. <http://cep.lse.ac.uk/seminarpapers/09-12-11-SB.pdf>
- Carayol N, Roux P (2007) The Strategic Formation of Inter-Individual Collaboration Networks: Empirical Evidence from Co-invention Patterns. *Annales d'Economie et de Statistiques* 87-88: 275-301
- Crespo J, Suire R, Vicente J (2013) Lock-in or lock-out? How structural properties of knowledge networks affect regional resilience. *Journal of Economic Geography*, forthcoming. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2034901

- Cranmer SJ and Desmarais BA (2011) Inferential Network Analysis with Exponential Random Graph Models. *Political Analysis* 19(1): 66-86
- Daraganova G et al (2012) Networks and Geography: Modelling Community Network Structures as the Outcome of Both Spatial and Network Processes. *Social Networks* 34(1):6-17
- Erdős P, Renyi A (1959) On random graphs. I. *Publicationes Mathematicae Debrecen* 6: 290-297
- Feldman MP (1993) An Examination of the Geography of Innovation. *Industrial and Corporate Change* 2(3)
- Feldman MP, Florida R (1994) The Geographic Sources of Innovation: Technological Infrastructure and Product Innovation in the United States. *Annals of the Association of American Geographers* 84(2): 210-229
- Frachisse D (2010) L'analyse des réseaux sociaux appliquée aux PCRD: Enjeux méthodologiques et outils pour l'évaluation. Dissertation, CREUSET, Université de Saint-Etienne.
- Frank O, Strauss D (1986) Markov Graphs. *Journal of American Statistical Association* 81(395): 832-842
- Gay B, Dousset B (2005) Innovation and network structural dynamics: Study of the alliance network of a major sector of the biotechnology industry. *Research Policy* 34: 1457-1475
- Geuna A (1998) Determinants of University Participation in EU-funded R&D Cooperative Projects. *Research Policy* 26: 677-687
- Goyal S, van der Leij MJ, Moraga-González JL (2006) Economics: An Emerging Small World. *Journal of Political Economy* 114(2): 403-412
- Granovetter M (1985) Economic Action and Social Structure: The Problem of Embeddedness. *The American Journal of Sociology* 91(3): 481-510
- Hanneke S, Fu W, Xing EP (2010) Discrete Temporal Models of Social Network. *Electronic Journal of Statistics* 4: 585-605
- Hoekman J, Frenken K, van Oort F (2009) The Geography of Collaborative Knowledge Production in Europe. *The Annals of Regional Science* 43(3): 721-738
- Hoekman J, Frenken K, Tijssen RJW (2010) Research Collaboration at a Distance: Changing Spatial Patterns of Scientific Collaboration within Europe. *Research Policy* 39: 662 – 673
- Jackson MO, Wolinsky A (1996) A strategic Model of Social and Economic Networks. *Journal of Economic Theory* 71: 44-74
- Johnson C, Gilles R (2000) Spatial social networks. *Review of Economic Design* 5: 273-299
- Krivitsky PN (2012) Exponential-family Random Graph Models for Valued Networks. *Electron J. Statistics* 6: 1100-1128.
- Krivitsky PN, Handcock MS (2010) A Separable Model for Dynamic Networks. arXiv:1011.1937v1
- Maggioni MA, Nosvelli M. and Uberti TE (2007) Space vs. Networks in the Geography of Innovation : A European Analysis. *Papers in Regional Science* 86: 471-493
- Mairesse J and Turner L (2005) Measurement and Explanation of the Intensity of Co-publication in Scientific Research: An analysis at the Laboratory Level. NBER Working Paper 11172
- Massard N, Mehier C (2009) Proximity and Innovation through an "Accessibility to Knowledge" Lens. *Regional Studies* 43 (1)
- Newman MEJ (2001) Clustering and Preferential Attachment in Growing Networks. *Physical Review E* 64(2)
- Nowicki K and Snijders T (2001) Estimation and prediction for stochastic blockstructures. *Journal of the American Statistical Association* 96: 1077-1087
- Paier M, Scherngell T (2008) Determinants of collaboration in European R&D networks: Empirical evidence from a binary choice model perspective. NEMO Working Paper #10
- Park J, Newman MEJ (2004) The Statistical Mechanics of Networks. *Physical Review E* 70(6).
- Pattison P, Robins G (2002) Neighbourhood-based models for social networks. *Sociological Methodology* 32: 301-337
- Polanyi M (1966) *The Tacit Dimension*. Doubleday & Company Inc., Garden City, New York.
- Ponds R, van Oort F, Frenken K (2007) The Geographical and Institutional Proximity of Research Collaboration. *Papers in Regional Science* 86(3): 423-443

- Powell WW, Koput KW, Smith-Doerr L (1996) Inter-organizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly* 41(1): 116-145
- Powell WW et al (2005) Network Dynamics and Field Evolution: The Growth of Inter-organizational Collaboration in the Life Sciences. *American Journal of Sociology* 110(4): 1132-1205
- Robins G et al (2007) An Introduction to Exponential Random Graph (p*) Models for Social Networks. *Social Networks* 29: 173-191
- Scherngell T, Barber M (2009) Spatial interaction modelling of cross-region R&D collaborations: empirical evidence from the 5th EU framework programme. *Papers in Regional Science* 88(3)
- Scherngell T and Lata R (2011) Towards an Integrated European Research Area? Findings from Eigenvector Spatially Filtered Spatial Interaction Models Using European Framework Programme Data, *Papers in Regional Science*. doi: 10.1111/j.1435-5957.2012.00419.x
- Snijders T et al (2010) Introduction to Stochastic Actor-based Models for Network Dynamics. *Social Networks* 3: 44-60
- Steglich C et al (2010) Dynamic Networks and Behaviour: Separating Selection from Influence. *Sociological Methodology* 40(1): 329-393
- Ter Wal ALJ (2013) The dynamics of the inventor network in German biotechnology: geographical proximity versus triadic closure. *Journal of Economic Geography* (2013) doi: 10.1093/jeg/lbs063.
- Vinciguerra S, Frenken K, Valente M (2010) The Geography of Internet Infrastructure: An Evolutionary Simulation Approach Based on Preferential Attachment. *Papers in Evolutionary Economic Geography* 10.06, Utrecht University Urban and Regional Research Centre.
- Wasserman S, Pattison P (1996) Logit Models and Logistic Regression for social Networks: I. An Introduction to Markov Graphs and p*. *Psychometrika* 61(3): 401-425