CIRCLE

LUND UNIVERSITY

Paper no. 2014/11

Social Network Analysis Methodologies for the Evaluation of Cluster Development Programs

Elisa Giuliani (elisagiuliani@gmail.com)
Department of Economics and Management, University of Pisa
CIRCLE, Lund University

Carlo Pietrobelli (carlop@iadb.org)
Inter-American Development Bank

This document was previously published by the Inter-American Development Bank with reference number IDB-TN-317.

This version: June 2014

Centre for Innovation, Research and Competence in the Learning Economy (CIRCLE)
Lund University
P.O. Box 117, Sölvegatan 16, S-221 00 Lund, SWEDEN
http://www.circle.lu.se/publications

WP 2014/11

Social Network Analysis Methodologies for the Evaluation of Cluster Development Programs

Elisa Giuliani and Carlo Pietrobelli

Abstract

Cluster development programs (CDPs) have been adopted widely in many countries worldwide. Many such programs aim to promote economic development by forming and strengthening inter-organizational networks. Despite their widespread diffusion, we know very little about CDP outputs or the impact CDPs have on host regions and their populations. Evaluation studies are beginning to appear, but the overall concern is that a distinct evaluation concept and method with a focus on CDPs is not yet available. The objective of this paper is to address this limitation, by proposing a novel methodological approach in the evaluation of CDPs based on the application of concepts and methods of social network analysis (SNA).

JEL codes: O22, O29, Z13

Keywords: cluster development programs, policy evaluation, social network analysis

Social Network Analysis Methodologies for the Evaluation of Cluster Development Programs

Elisa Giuliani

Via C. Ridolfi, Facoltà di Economia, University of Pisa, Italy
Phone: +39 050 2216280; Email: elisagiuliani@gmail.com
&

CIRCLE, Lund University

Carlo Pietrobelli

Inter-American Development Bank, Washington, D.C., and
CREI University of Roma Tre, Italy
Phone: +1 202 623 2423; Email: carlop@iadb.org

Abstract

Cluster development programs (CDPs) have been adopted widely in many countries worldwide. Many such programs aim to promote economic development by forming and strengthening inter-organizational networks. Despite their widespread diffusion, we know very little about CDP outputs or the impact CDPs have on host regions and their populations. Evaluation studies are beginning to appear, but the overall concern is that a distinct evaluation concept and method with a focus on CDPs is not yet available. The objective of this paper is to address this limitation, by proposing a novel methodological approach in the evaluation of CDPs based on the application of concepts and methods of social network analysis (SNA).

JEL Classification: O22, O29, Z13
Keywords: cluster development programs, policy evaluation, social network analysis

Acknowledgements

The authors acknowledge the discussions from the participants and wish to thank Elias Carayannis, Gabriel Casaburi, Rafael Corredoira, Claudio Cortellese, Gustavo Crespi, Emiliano Duch, Eduardo Fernandez Arias, Jordi Jaumandreu, Natalia Laguyas, Alessandro Maffioli, Manuel Pacheco, Ernesto Stein, and Rodolfo Stucchi for their comments. This paper is an output of the RG-T1861 "Policies and Institutions for Productivity in Latin America and the Caribbean."

Disclaimer: This paper was previously published as Inter-American Development Bank No. IDB-TN-317.

Section 1. Introduction

Cluster policies and other regional-level policy initiatives aim to promote local collective processes of growth and innovation and have become popular in many countries worldwide. Cluster development programs (CDPs) became fashionable during the 1990s, and a recent survey identified more than 1,400 currently ongoing projects in the world (Ketels, Lindqvist, and Sölvell, 2006). Such programs have been promoted by governments in most European Union (EU) countries (Lundquist and Power, 2002; European Commission, 2002a); in the United States (Hospers, Desrochers, and Sautet, 2009); in Japan (Ganne and Leclerc, 2009; Mori, Kajikawa, and Sakata, 2010); and, often upon the suggestion of donors and international consultants, in many developing and emerging economies, including Latin America and Caribbean (LAC) regions (Condo and Monge, 2002; Peres, 2004; Ketels *et al.*, 2006; MIF, 2007; Pietrobelli, van Oyen, and Dobinger, 2009; Ferraro, 2010; Pacheco, 2010).

There is no single definition of "cluster policy" or "cluster development program," which is plausibly due to the concept of a "cluster" being subject to multiple interpretations (Martin and Sunley, 2003).[1] In this paper, we maintain that CDPs "refer to all those efforts of the government to develop and support clusters in a particular area." (Hospers *et al.*, 2009) They are often a combination of industrial and science and technology (S&T) policies, which are adopted at regional or local levels, in some cases by regional governments themselves. While there is a wide variety of CDPs (see below), they share at least three aspects on which international experts converge. First, they do not tend to target individual firms or sectors (as pure industrial policies), but rather a group or a network of actors that are typically localized in a geographically bounded area (whose boundaries may vary according to the context). Second, there is no "one-size-fits-all" CDP model, as these types of programs have to be crafted on the basis of local specificities and idiosyncrasies: the history, the stage of the cluster life cycle, and the specificities of the industry are all relevant aspects when deciding on the right initiative (Landabaso, 2000; Europe INNOVA, 2008). Third, all CDPs are grounded on the idea that growth and development are based on both the search for more efficiency in production and for higher innovation, both of which are best achieved when firms collaborate and share resources.

To accomplish their objectives, CDPs often involve different sets of initiatives, which may co-exist within the same policy intervention and be promoted by different actors (Raines, 2001;

[1] The cluster concept is derivative of the original industrial district concept developed by Alfred Marshall, who defined it as a concentration of "large numbers of small businesses of a similar kind in the same locality" (Marshall, 1920: 277). In a very influential work, Porter (1998) defines clusters as "geographic concentrations of interconnected companies, specialised suppliers, service providers, firms in related industries, and associated institutions (…) in particular fields that compete but also cooperate" (pp. 197–8). It is beyond the scope of this paper to provide a review of the different conceptions and definitions of clusters adopted through the years by scholars and policymakers. The interested reader can refer to Markusen (1996) and Martin and Sunley (2003).

European Commission, 2002b; Borrás and Tsagdis, 2008). Very often they are aimed at institutional building, and at pursuing the formation and strengthening of those public and/or private organizations that play a pivotal role in spurring innovation, technology transfer, training, export and internationalization, and marketing activities. A typical example is setting up training institutes for small and medium-sized enterprises (SMEs), such as the business support services in the Italian context (Pietrobelli and Rabellotti, 2007). In other cases, cluster policies revolve around providing new infrastructure facilities—a case in point being the development of science parks and incubators. Despite such a variety of initiatives, a central tenet of many cluster policies is the promotion of linkages and networks. In fact, CDPs are often about:

> "*stimulating the links* to the local business environment through public-private dialogues, defining *joint research needs*, *co-development* between contractors and suppliers and so on" (European Commission, 2002b; emphasis added)

and about:

> "boosting the development of a competitive private sector and contributing to poverty reduction by *building sustainable linkages* both among SMEs and between SMEs, large(r) scale enterprises and support institutions." (Pietrobelli *et al.*, 2009; emphasis added).

Policymakers attach great value to creating and strengthening networks. Their view is grounded on a well-consolidated body of academic research that shows that, in modern economies, firms that are embedded in systems of social relations enjoy a privileged position relative to isolated ones. In this respect, Granovetter (1985) suggests transaction costs can be kept to a minimum if firms are embedded in networks of social relations that monitor and sanction opportunistic behaviors and malfeasance (Gulati, 1995). We know that firms return to networks as these deliver certain advantages when compared with other governance structures (Powell, 1990). First, they permit the informal exchange of unique and idiosyncratic assets such as knowledge or know-how, which market mechanisms are unlikely to transact. Second, they are a relatively loose and rapid way in which to put individuals or organizations in contact, even when these are not formally connected to each other. Third, they have the power to maintain stable and high-quality relationships over time, fostering trust and reciprocity.

The advantages of networks are particularly striking for innovation, which is now widely recognized as being a social process involving the interaction, alliance, and cooperation of different actors (Freeman, 1991; Powell, Koput, and Smith-Doerr, 1996). In developing economies, where market failures and institutional weaknesses may be particularly severe, firms may find inter-organizational networks can be used as safety nets against uncertainty and unfavorable business climates. They use networks to access resources, reduce information asymmetries, enable higher bargaining power versus other market counterparts, strengthen their lobby power towards governments, and enable firms to upgrade their capabilities (Guillén, 2000; Khanna and Rivkin, 2001; Mesquita and Lazzarini, 2008; McDermott, Corredoira, and Kruse, 2009). In poor regions, networks are considered also to be an important piece of the poverty-alleviation puzzle. Better linkages may be a functional way to obtain efficiency gains that an individual firm, especially if micro and small, cannot attain (Schmitz, 1995). For rural farmers in Africa or Asia, networks may be the sole way through which a buoy or a donkey to reach the nearest village can be bought and/or shared. Hence, collaborative linkages allow poor entrepreneurs to pool resources and efforts to achieve shared economic goals (Pietrobelli et al., 2009). The recent hype on networks as a direct target of CDPs is based on these grounds.[2]

While CDPs have proliferated across the globe, their impact is unknown to most, and attempts to evaluate their impact are still in their early stages (Fromhold-Eisebith and Eisebith, 2005; Quiroz, 2007; Baruj, 2007; Nishimura and Okamuro, 2011). This is because of the relatively recent implementation of CDPs, and also because there are "intricate methodological complexities involved" (Schmiedeberg, 2010, p. 1251) in such an evaluation process, and therefore a "distinct evaluation concept or toolkit with focuses on cluster policy is not yet available" (p. 1251). The objective of this paper is to address these limitations by proposing a novel methodological approach for CDP impact evaluation based on the application of concepts and methods of social network analysis (SNA)—an approach that has hardly been used for policy evaluation purposes. This paper is directed to evaluators and cluster scholars with an interest in evaluation programs, and it is designed to be a primer on the use of SNA for cluster evaluation purposes. It is primarily directed to evaluators and scholars with no prior knowledge of SNA, but it may equally be a valuable resource for SNA experts who wish to apply this methodology to cluster evaluation. More specifically, Section 2 aims to illustrate the usefulness of SNA for CDP impact evaluation. Section 3 goes much deeper into the opportunities and constraints of the SNA methodology and offers a practical guideline for the application of this methodology. Section 4 concludes this paper.

Section 2: Why SNA Is Useful for CDP Evaluation

2.1 The Need to Improve the Measurement of Networks

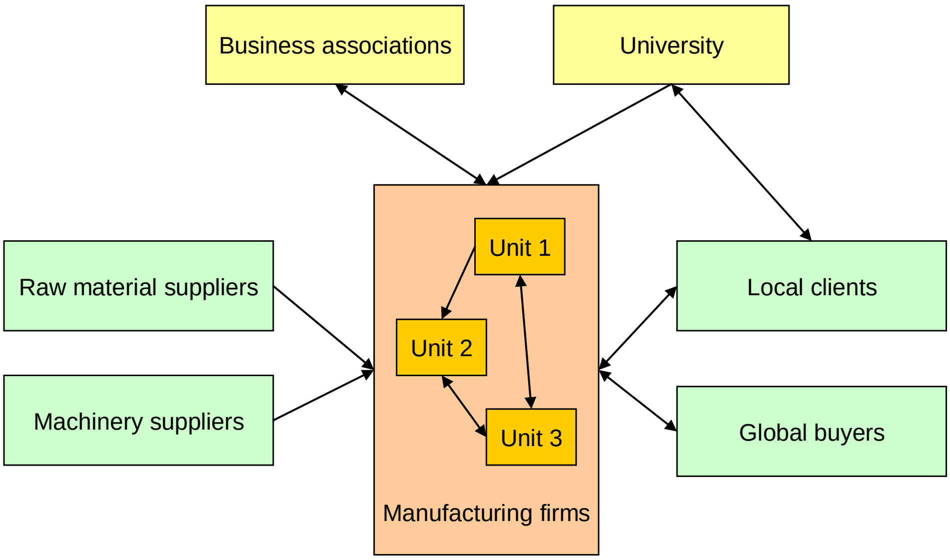
The aim of this section is to stress the importance of studying networks in more analytical terms than has been done in prior evaluation studies, so as to inform policymakers about CDP design and evaluation. On the one hand, policymakers have put great emphasis on networks as a means to

[2] Hereinafter, we will use the general expression "cluster development programs" (CDPs) to refer narrowly to those intended to foster networks.

stimulate learning and innovation and to achieve efficiency gains. Yet on the other hand, most of the available evaluation attempts are based on a very poor understanding of what networks are, and key concepts like “networking,” “connectivity,” “connections,” and “linkages” are often measured through rather loose and rough indicators. For instance, to account for the degree of connectivity of firms to their regional suppliers in the United Kingdom, McDonald et al. (2007) simply distinguish between “deep” and “shallow” regions based on the level of local connectedness as reflected in the input-output tables. Other studies simply rely on the perceptions of their respondents about whether given CDPs have stimulated the formation of collaborative activities (e.g., joint production, joint sales, joint research and development) (Huggins, 2001; Ketels et al., 2006) or use the number and frequency of formal or informal cooperation among cluster members (Raines, 2002; Rosenfeld, 2002; FOMIN, 2010; Nishimura and Okamuro, 2011). In other cases, the simple participation of a firm to a local business association is considered a networking process (Aragón et al., 2009). Furthermore, when it comes to mapping relationships, the most conventional approach is to identify flows between firms in one industry and suppliers in another, as well as between firms and other local private or public organizations (for an illustration, see Figure 1). However, as rightly suggested by Rosenfeld (2002), “most maps are very general, showing cluster members as boxes but with little precise knowledge of the strength of the linkages” (p. 17). More importantly, this type of network mapping only captures the linkages between general categories of actors (e.g., “suppliers of raw materials” or “local clients”); it fails to account for the vast heterogeneity existing within each category, with some actors playing much more critical roles in shaping the network than actually acknowledged in such maps. Hence, while current approaches to measure networks are plausible and justifiable by practical motivations, as they offer a simple and rather inexpensive way of accounting for the presence of networks, we contend here that we should be cautious about their use in CDP evaluations, as they may hide more than they reveal.[3]

[3] This approach to the analysis of networks is not just a characteristic of CDP evaluations. It is also a dominant approach in the research on local innovation systems.

Figure 1 Example of Conventional Mapping of Networks in Clusters

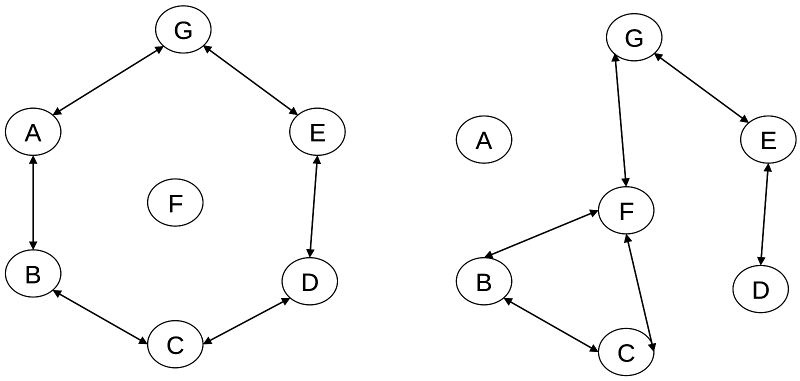


Source: Authors' elaboration.

To explain the relevance of studying networks in greater depth than is usually done by conventional approaches, we propose a simple comparative example of two networks (see Figure 2). A network is defined as "a finite set or sets of actors and the relation or relations defined on them" (Wasserman and Faust, 1994). The actors of the network can be of a different nature (individual entrepreneurs or firms, public organizations, etc.), while the link represents *one* type of relationship existing between the different actors.

Suppose that in Figure 2 the actors indicated with A, B, C, and so on, are firms (firm A, firm B, firm C, etc.), and the ties represent the flow of some kind of asset (e.g., advice). The structure of any of these two networks is the result of the connectivity choices of firm A, firm B, firm C, and so on. By measuring the average number of ties established by each firm, as well as the density of ties in the networks, it is evident that these indicators are roughly the same for both networks. However, by examining the way linkages are distributed in the topological space, it is clear that these two networks remarkably differ. The network illustrated in Figure 2a is completely a-hierarchical with one disconnected actor (F). The second network (Figure 2-b) is instead a hierarchical network, where F becomes a central node, a bridging actor between most of the network's actors, except for A, which is isolated. As explained in greater depth in the sections that follow, these differences are very important, as they have implications on the way assets (e.g., advice, goods, resources, etc.) are circulated and shared: much more evenly in the case of the first network, where no firm seems to occupy a dominant position, than in the case of the second network, where F sits in a highly strategic and powerful position. Through standard methodological approaches it is virtually impossible—especially with larger networks—to spot these structural differences.

Figure 2 Two Different Network Structures



(a) **(b)**

Source: Authors' elaboration.

A different method is required, and SNA proves particularly valuable in network measurement.[4] SNA is a distinct research perspective within the social sciences, and it is based on the assumption that relationships among interacting actors are important to explain their nature, behavior, and outputs. To rigorously study relationships, SNA uses graph theory, a mathematical discipline that is considered to have been initiated in the 18th century (Newman, 2003), and to have been applied to social science at the beginning of the 20th century—with the social psychologist J. L. Moreno being one of the key precursors (Moreno, 1934). Based on the application of graph theory, SNA is considered akin to an "organizational X-ray" tool (Serrat, 2009), as it makes visible what for other methodological approaches is invisible—as illustrated in the example above. But what is the value added of this type of analysis? Why is it so important that relations become visible? What do we learn from unravelling the structure of a network and the position of an actor within a network? The following sections will attempt to answer these questions.

2.2 The Advantages of Analyzing Networks

2.2.1 The Importance of Being Well Positioned

As illustrated in Figure 2, not all actors are equally positioned within a network. Actor F is positioned differently in Figure 2-a than in Figure 2-b. Such a difference is a rather critical aspect to detect: depending on the nature and characteristics of the linkages, an actor's position may reflect its power, prestige, or access to or control of resources. Central actors are generally considered to be

[4] An entire special issue on *New Directions for Evaluation* (107, Fall, 2005) was devoted to the adoption of SNA for policy evaluation purposes. See also Davis (2003) and Section 3 for existing applications to cluster policies.

in advantageous positions (Laumann and Pappi, 1976; Freeman, 1979).[5] For instance, in a communications network, central firms may be better positioned in accessing information: the higher the number of direct ties an actor has with others in the network, the higher the actor's opportunities for learning and accumulating experience and skills. Also, firms with multiple information sources are considered less likely to "miss vital information" (Bell, 2005). However, too many linkages may overload an actor, as building connections involves an important opportunity cost in terms of time invested in forming or maintaining a relationship that may be used for alternative activities.

The number of direct ties an actor holds with others in the network—technically, the degree centrality—is one of the most basic and intuitive ways to measure centrality (Table 1-a). However, depending on the nature of the ties, and on the type of impact or output the actors are seeking, other types of less intuitive centralities may be more relevant (Freeman, 1979; Borgatti and Everett, 2006) (see Table 1 for an overview). For instance, in an influential paper, Bonanich (1987) suggests that, in bargaining situations, power comes from being connected to those who are powerless, as being connected to powerful others who have many potential trading partners reduces one's bargaining power. Hence, the power of an actor may be tied to the many direct ties of that actor as well as to the little ties of its direct contacts. On the contrary, in other types of networks (e.g., flow of technical knowledge), it may be advantageous to be tied to actors who have many connections, as this guarantees access to an even higher number of knowledge sources (Table 1-b). Research shows that the impact that this type of centrality has on innovative performance follows an inverted U-shape pattern, because central actors' positions will receive large amounts of information that, beyond a certain level, would overload and overwhelm them and taper off their ability to generate quality knowledge (Paruchuri, 2010).

In other circumstances, the advantages of being central may stand in the control that an actor has over the flow of goods, people or other material or immaterial assets. In this case, a central and powerful actor is one that bridges connections between other actors which would otherwise not be connected: "actors on whom others are locally dependent [for getting access to assets and resources] are central in the network" (Wasserman and Faust, 1994). Network scholars refer to these actors as having high "betweenness centrality" (Table 1-c); their power is related to being considered essential to the network, as their removal would produce a disruptive effect. An actor that is the only (or just one of the few) channel through which other actors can get connected can exert a certain power on its direct ties, and even influence negatively their operations. For instance, in

[5] The idea of centrality as applied to human communication was introduced by Bavelas and colleagues at M.I.T. in the late 1940s. He was specifically concerned with communication in small groups, and he hypothesized a relationship between structural centrality and influence in group processes.

market relationships, a firm playing a brokering role between many SMEs and a large global buyer may privilege the relationship with the buyer, thus accepting lower prices or unfavorable market conditions, to keep this tie. Meanwhile, the broker could transfer those constraints to its small suppliers, which in turn may be "forced" to accept bad deals (e.g., tight delivery schedules or stringent cost requirements) to maintain the tie with the broker (for an application on the staffing sector, see also Fernandez-Mateo, 2007).

This example shows how much an actor's advantage may depend on the degree to which an actor's direct contacts (alters) are (or are not) connected to each other. A further classical distinction is made between the case in which an actor is positioned in a network where its alters are all densely connected to each other (high closeness, see Table 1-d), and the case in which an actor sits on a structural hole, with all or most of its alters being unconnected to each other (Table 1-e). These two positions convey different types of advantages. High closeness is normally considered a precondition for the emergence of trustful relations—an important governance mechanism, since it reduces both uncertainty and information asymmetries in the interactions between two actors (Coleman, 1988). Also, close ties typically allow the exchange of more fine-grained information, which is more proprietary and tacit than the information exchanged in open networks; therefore, close ties also entail effective joint problem-solving arrangements that speed up responses to the market (Uzzi, 1997). Research has shown that, in innovation or communications networks, the higher the degree of closure of an actor's (ego-centered) network, the more innovative will be the actor, as this helps firms achieve deep understanding of a specific innovation (Zaheer and Bell, 2005). However, similarly to other types of centralities, when firms are too closely embedded in a network, the risk is that they get "trapped in their own net" (Gargiulo and Benassi, 2000). In fact, close ego-centered networks may breed relational inertia and obligations for reciprocity. In turn, this may have the effect of cementing relationships into a stable network structure, even when these relationships are no longer beneficial and result in firms relying only on knowledge from their trusted alters, generating a risk of negative technological lock-in (Grabher, 1993) and hampering innovation performance (Giuliani, 2008).

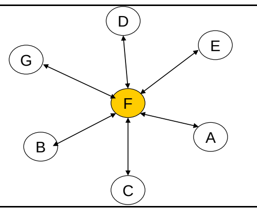
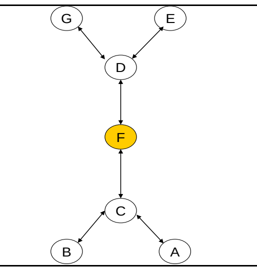
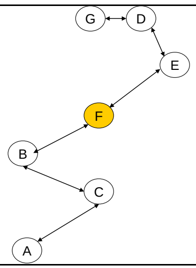
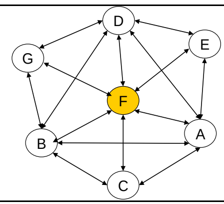
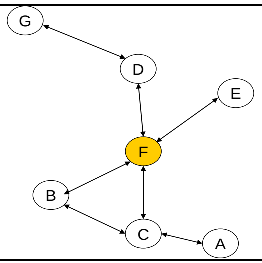
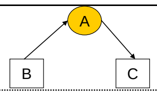
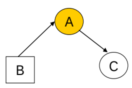
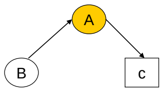
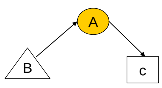
In some cases, scholars have argued that radical innovations and true creative ideas are better reached through the search for informational "diversity" (Laursen and Salter, 2006). Drawing on Burt's structural holes theory (1992; 2001), network scholars have suggested that such diversity is best achieved when an actor's direct contacts are *not* densely connected to each other, and thus there is a "hole" in the knowledge network structure (Table 1-e). Structural holes theory suggests that firms acting as brokers in a network have access to potentially more diverse knowledge, which enhances the exploitation of new ideas and the promotion of radical innovations (Ahuja, 2000;

Rowley, Behrens, and Krackhardt, 2000; McEvily and Zaheer, 1999; Zaheer and Bell, 2005). Furthermore, actors situated on structural holes economize the number of ties required to access unique information and can earn control benefits because they act as brokers between disconnected partners—an advantage that is similar to that obtained by actors with the high betweenness centrality mentioned earlier (Baum, Shipilov, and Rowley, 2003).

So far in this report, we have discussed network positions considering equivalent actors. However, it is also possible that a network is composed of actors that belong to non-overlapping communities. Thus, for instance, a network may be formed by firms, but also by different university departments and business associations, and an advantageous strategic position would be that of being at the interface of different communities. In their seminal paper, Gould and Fernandez (1989) identify different types of brokerage roles, depending on the types of communities an actor is able to connect. This may be the case of the *itinerant broker*, who connects actors that have the same affiliation but whose affiliation is different from that of the broker (e.g., a firm connecting two different universities); the *gatekeeper/representative*, who connects an actor having the same affiliation as the broker with an actor of a different affiliation (e.g., a firm connecting another firm with a university); and the *liaison*, who connects actors that have different affiliations from the broker and from each other (e.g., a firm making the connection between a university and a business association) (see Table 1-f for an illustration). Actors connecting different communities have access to resources that are enriching and can be vital for the whole community. In a study on Chile and South Africa, Giuliani and Rabellotti (2011) show that the most talented university researchers are more likely than others to act as brokers between the local industry and their international colleagues in the academia.

The extent to which different brokering roles matter, or lead to an improvement of the conditions of the broker itself or of the other actors in their community, will depend very much on the nature of linkages and on other contextual factors.

Table 1 Examples of Network Positions, Beneficial Effects, and Limits

SNA concept	Brief description	Illustration	Advantages/benefits	Limits
(a) Degree centrality	Number of direct ties an actor has with others in the network		Easy access to information, knowledge, and any type of resource	Too many connections can be time-consuming, not always rewarding
(b) Bonacich centrality	Centrality of an actor dependent on the centrality of its direct contacts (alters)		Power (if alters have low centrality), access to resources (if alters have high centrality)	Too many connections may overload the actor
(c) Betweenness centrality (see also "Structural holes")	Degree to which an actor is able to connect others that will be otherwise disconnected		Gatekeeping, influence, dependence, control	If there are only a few actors with high betwenness centrality, they may easily disrupt the network (vulnerability risk)
(d) Closed ties	High local connectivity between an actor's alters		High trust, high-quality knowledge, joint problem-solving, reduction of transaction costs	Too much closure is detrimental and leads to lock-in
(e) Structural holes (see also "Betweenness centrality")	When an actor's alters are not/are poorly connected to each other		High level of knowledge diversity, high opportunities for creativity and radical innovations, efficiency and control in ties	Does not have the advantages of network closure
(f) Brokerage roles	*Itinerant broker*		It is possible to identify the degree to which an actor plays any of these roles. Actors connecting different communities or subgroups (signalled by different node shapes in the figure) have access to resources that are different, and they can also exert control on the actors that they are connecting. The advantages and limits of any of these roles depend very much on the nature of linkages and context.	
	Gatekeeper			
	Representative			
	Liaison			

2.2.2 How Linkages Are Structured into a Network

We have seen that network positions hold certain advantages and constraints for individual actors. However, CDP practitioners often aim to promote the growth of a *whole* geographical area or of a *whole* community of firms and entrepreneurs, rather than that of individual firms. To move from the actor to the network level, we need to shift our perspective and focus on the overall network structure — in other words, on the way the linkages between the actors are distributed. Newman (2003) tells us that "real networks are non-random in some revealing ways that suggest both possible mechanisms that could be guiding network formation, and possible ways in which we could exploit network structure to achieve certain aims." (p. 180). The non-random distribution of most networks means that their structure is due to their actors' connectivity choices, which reflect the actors' own strategies, purposeful and selective choices, and bounded rationality.

Connectivity choices by individual actors tend to be somewhat myopic, as actors most likely know why they need a connection and have an idea as to whom they want to be connected, but they have no vision of how their choices (and that of others forming linkages at the same time) shape the overall network structure. Depending on an actor's network position, even a single change in a tie may lead to significant modifications in the overall network structure. As suggested by Watts (2004), "large changes in the structure of a system could be derived by even subtle modification in the network structure and modifications may be imperceptible to actors with only local knowledge of the network" (p. 246). While individual actors may be unaware of the global consequences of their local connectivity choices, it remains of great interest for analysts of cluster policies to have a global picture of a network structure so as to foresee which actors are most likely to generate severe, devastating, or disruptive effects on the network.[6] Furthermore, the structure of a network can have significant implications for the collective dynamics of a system, whose connectivity the network represents (Watts, 2004). Different network structures convey different types of collective advantages and disadvantages and can then influence the growth trajectories of regions and clusters. Such differences are generally overlooked by CDP analysts, for whom a well-functioning network often corresponds to one with dense connectivity, while networks with sparse linkages are often dismissed as weak and not functioning (for a critical discussion on this matter, see Staber, 2001).

Given the non-random structure of networks, one of the major concerns of social network analysts is identifying subgroups of actors that display higher average connectivity than the rest of the actors in the network. Scholars typically refer to these groups as "cohesive" subgroups, which

[6] In the context of networks, the expression "local" refers to actors that have close or direct connections, and "distant" refers to actors that are not reachable directly or with few steps. Similarly, the expression "global" refers to the overall structure of the network. These expressions do not therefore refer to the geographical distance of actors in space, but only to the topological distance, measured by the number of ties or steps separating actors.

means a "subset of actors among whom there are relatively strong, direct, intense, frequent or positive ties" (Wasserman and Faust, 1994). In general terms, cohesive subgroups are characterized by mutual and frequent ties, the closeness or reachability of its members, and a high frequency of ties, yet these subgroups can be formalized in different ways, depending on the properties of the ties among subsets of actors. Hence, there are different types of subgroups within a network, leading to differences in the overall network structure. More importantly, the characteristics of different subgroups are associated with different types of benefits or disadvantages for network members. The earliest research on subgroups focused on the generation of cliquish friendship networks (Luce and Perry, 1949). "Cliques" are groups of *at least* three actors that are all connected to each other.[7] A giant clique could involve hundreds of actors, all connected to each other (Table 2-a). Provided that the nature of ties has some valuable content, cliquish networks have the advantage of guaranteeing a cooperative environment, whereby social monitoring, trust, and resource-sharing are bound to be high.[8] Also, cliquish networks are by definition a-hierarchical places where resources are distributed in a highly egalitarian way. In reality, however, very few networks are fully cliquish. Most networks are fragmented and display properties that depart from the densely connected structure of a clique: often, networks are formed by many smaller non-overlapping cliquish structures, which can be entangled with each other in very different ways.

In small networks, cliques can be connected to each other by sparse or weak ties (Table 2-b-i). In larger networks, this structural feature became known as "small world" (Table 2-b-ii). Building on the famous small world experiment conducted by Stanley Milgram in the late 1960s at Harvard, Watts and Strogatz (1998) developed a formal mathematical model describing systems with small world properties. These systems have two core properties: 1) high local density, which means that actors have dense connections with their neighbors (local cliques); and 2) few connections with other distant actors (clique-spanning ties). Small world networks have been well documented in sociological and economic literature (firm co-ownership networks by Kogut and Walker, 2001; the musical industry by Uzzi, Spiro, and Delis, 2002; collaboration networks of board of directors by Davis, Yoo, and Baker, 2003; scientist networks by Newman, 2001). Their advantage is that they are efficient network structures; despite the overall relatively low density of ties, as a result of local cliquishness and clique-spanning ties, actors in a small world network are linked with each other through a relatively small number of intermediaries. The importance of small world structures in inter-organization systems stems from "their great efficiency in moving

[7] This definition can be relaxed. There are different notations that measure clique-like structures, although using a less strict definition (e.g., *n*-cliques, *n*-plexes; see Section 3).

[8] We assume that the ties between two actors are "positive" ties, and thus serve to channel resources and other assets. In the opposite case—i.e., a tie representing a criminal linkage or the flow of something that can damage the parties—the discussion is different.

information, innovations, routines, experience, and other resources that enable organizational learning, adaptation and competitive advantage" (Baum *et al.*, 2003). Furthermore, this structure has the advantage of benefiting from a high level of local trust, cooperative environment, shared consensus, and mental models through the high density of local cliques. At the same time, it also guarantees that local cliques do not remain isolated, as some members are also connected with distant actors. Although in some cases, distant connections may be due to chance (Watts, 2004), Baum *et al.* (2003) suggest that, in business organizations, distant ties may be the result of strategic manoeuvring by local firms, some of which deliberately search for a competitive advantage by spanning beyond local ties.

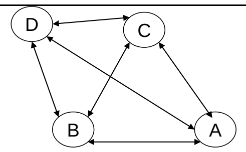
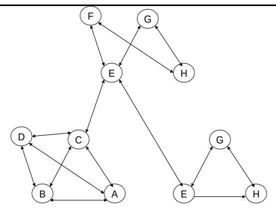
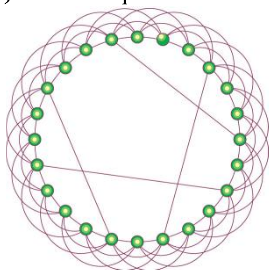
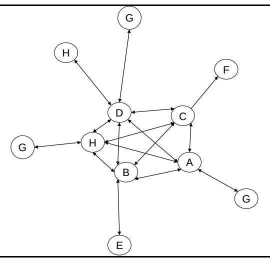
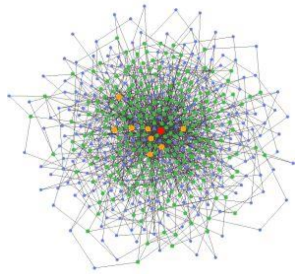
Some networks are efficient and hierarchical. A typical hierarchical structure is the "core-periphery" (Table 2-c). As the name indicates, a core-periphery structure is composed of a densely connected core (a cliquish-like subgroup) and a set of "angers-on" actors (i.e., the periphery), which are loosely connected to the core and very poorly connected to each other (Borgatti and Everett, 1999). If the nature of ties permits it, actors in the core often constitute an elite and benefit from the advantage of being part of a central group. In a paper on wine clusters in Chile, Giuliani and Bell (2005) show that only firms with high absorptive capacity were able to be part of the core, while the peripheral firms were only marginally included in the local knowledge-generating networks—a position that was seen as hampering their learning and innovation opportunities. In a completely different context, such as the Hollywood motion picture industry, Cattani and Ferriani (2008) show that individuals who occupy an intermediate position between the core and the periphery are in a favorable position to achieve creative results, while individuals who are either highly peripheral or over-entrenched in the social network of the core tend to be less creative. Hence, different positioning within a core-periphery structure leads to different benefits, with peripheral firms often being in a disadvantaged position. At the network level, this structure maintains a hierarchy in relationships that may generate and endure a divide between network actors.

An even more pronounced hierarchical structure, which has recently been found to characterize many real-world networks, became to be known as "scale-free" (Barabasi and Albert, 1999). Scale-free means that the distribution of the number of direct contacts (degree centrality distribution) an actor has is typically right-skewed with a heavy tail, meaning that a majority of nodes have a less-than-average degree of connection and that a small fraction of hubs are many times better connected than average (Table 2-d). This type of structure is generally considered to be the result of two concurrent mechanisms: population growth and preferential attachment (de Solla Price, 1976). Real networks grow in time as new members join the population. The mechanism of preferential attachment expresses the notion that newly arriving nodes will tend to connect to already well-connected nodes rather than poorly connected nodes. This is because new entrants typically suffer from a lack of information about whom to connect to. Rigorous quality judgments may be extremely costly to make (Gould, 2002), so new entrants into a community may prefer to connect to highly reputable actors. As reputation is socially derived (Podolny, 1993; Ibarra and Andrews, 1993; Lazega *et al.*, 2010), actors that, at a given point in time, will have accumulated a critical mass of linkages to guarantee high status and reputation, will become like honey to bears; they will be targeted by most new entrants in the network, thus reinforcing their centrality over time. These networks are hierarchical, and hierarchy in the context of industrial clusters may be easily associated with an uneven and high concentration of resources, as well as with a high degree of vulnerability. Brought into the industrial cluster context, scale-free networks are not too dissimilar to Markusen's (1996) "hub-and-spoke" districts, where the business structure is dominated by one or several large, vertically integrated firms surrounded by suppliers. Nor it is too dissimilar to accounts of clusters dominated by a lead firm or a large global buyer (Giuliani *et al.*, 2005), which orchestrate the local value chain, coordinate and arrange resources, and, in doing so, reduce the complexity that is inherent in any production site with high inter-firm division of labor.

2.3 Opportunities for Applying SNA to CDP Evaluation

Evaluation is undertaken with at least two key motivations: the first is to assess the effects of a policy to legitimate further policies in that direction; the second is to assess the learning process that is generated—by understanding what factors or mechanisms are held responsible for a policy's success, practitioners can learn how to make future projects more effective. Different qualitative and quantitative approaches exist to assess the effects of CDPs (for a recent review, see Schmiedeberg, 2010). Qualitative methodologies (e.g., participatory approaches, case studies, focus groups) are considered appropriate for process evaluation and understanding how some expected outputs have been generated (e.g., increased institutional capacity, well-performing networks). Also, qualitative methods can provide critical insights into beneficiaries' perspectives about the dynamics of a particular policy initiative, or the reasons behind its lack of results (Diez, 2001). While rich in qualitative details, these approaches generally fail to assess impact. Impact evaluation seeks to determine whether a policy had the desired effects on individuals, firms, and institutions (e.g., higher income, improved performance, higher innovation rates) and whether those effects are attributable to the policy. Hence, impact evaluation seeks the *causal* link between the policy and the impact—a link that can be found only through econometric analysis involving control groups (see Winters, Salazar, and Maffioli, 2010; Maffioli, Rodriguez Mosquera, and Stucchi, 2011; Nishimura and Okamuro, 2011).

Table 2. Examples of Network Structures, Advantages, and Limitations

SNA concept	Brief description	Illustration	Advantages/benefits	Limits
(a) A single cohesive set (clique)	A dense network where (almost) all actors are connected to each other		High level of trust, cooperation, support, and social monitoring	Redundant linkages, high opportunity costs, risk of "getting trapped in their own net"
(b) Small worlds	Non-overlapping cliques (high local closeness), connected by a few ties with distant actors	(i) Small cliquish structure (ii) Small worlds	Efficient structure, local dense ties (trust and cooperation), and distant ties (competitive advantage, search for diversity)	Success is dependent on actors with local and distant ties
(c) Core-periphery	A core of densely connected firms and a periphery with a few connections to the core and little intra-periphery ties		Core actors, as well as actors connecting the core to the periphery, may have advantages	Hierarchical structure, peripheral actors may suffer exclusion, uneven network structure
(d) Scale-free	Few hub firms holding all the connections, orchestrating a network		Hierarchical and organized management of the value chain	Very uneven structure, polarization of power and resources in a few actors, vulnerable to attacks to hubs

Given the inherent trade-offs between qualitative and quantitative approaches, there is a growing awareness about the benefits of integrated approaches to impact evaluation (Baker, 2000; FOMIN, 2011). SNA permits this integration. As acknowledged by scholars before us, it allows for a more rigorous assessment of CDP outputs based on the analysis of networks compared with conventional qualitative approaches. Among its numerous applications (see Section 3 for a full discussion), SNA can compare network structures prior to and after the policy treatment, and can observe whether significant changes in actors’ network positions have occurred during the policy implementation period. Moreover, SNA allows policymakers to assess whether the policy-targeted network(s) have acquired the structural properties expected by policy design (see Box 1 for stylized examples). For these reasons, there is a growing consensus around using SNA as a methodological tool for CDP impact evaluation: “as cluster policy explicitly focuses on the interaction of players, network analysis seems to be an adequate tool, at least for exploratory analysis on cluster development” (Schmiedeberg, 2010). This is reflected in recent CDP evaluation studies, which have adopted SNA as a visual, descriptive, or exploratory tool (see Table 3 for an overview). However, Schmiedeberg (2010) warns that “the value of a cluster is determined not only by network size and strength, but by the economic value the firms draw from these relationships (…). Thus, by *relying only on network analysis, the evaluation remains limited to intermediate outputs, but cannot draw conclusions on the real economic benefits* of the policy” (p. 401, emphasis added).

Table 3. Applications of SNA in the Evaluation of CDPs and Related Policies

Authors	Topic	Context	Use of SNA	Data	Method	Cluster policy?
Maffioli, 2005	Analysis of the impact of networking policies (PROFOS) on firm-level performance	PROFOS, Chile	Use of "rough" measures of centrality and network density	Firm-level survey	SNA and econometrics	No (Networking policy)
Matta and Donadi, 2007; Matta, 2010	Analysis of the evolution of inter-firm networks	Cordoba Province, Argentina	Descriptive analysis of different types of networks (density, structure, centrality)	Survey and project's primary data	SNA	Yes
Bellandi and Caloffi, 2009	Assessment of whether selected regional innovation policies promote the creation of well-functioning networks	Tuscany, Italy	Rudiments of SNA	Cooperation based on participation to RPIA-ITT projects; INTERREG funding	Rudiments of SNA	Yes
Russo and Rossi, 2009	Assessment of whether selected regional innovation policies promote the creation of well-functioning networks	Tuscany, Italy	Descriptive analysis of the network (structure)	Cooperation based on participation to RPIA-ITT projects	SNA	Yes
Ubfal and Maffioli, 2010	Impact of funding on research collaboration	Argentina	Measure of different types of researchers' centralities	Co-authorship data (ISI) Data on funding	SNA and econometrics (DID)	No (S&T policy)
Mori, Kajikawa, and Sakata, 2010	Impact assessment of regional cluster policies	Kanto and Koshinetsu region, Japan	Descriptive analysis of the network (structure)	Trade data between firms in the region	SNA	Yes

In this paper we argue the opposite; precisely, that SNA should not be limited to descriptive and exploratory evaluation exercises. SNA has a very important role to play in impact assessment analysis, as it generates highly valuable quantitative network indicators both at the level of the firm (or other relevant unit of analysis) and at the cluster level, which can be used in econometric estimates of impact assessment. Including firm-level centrality indicators in econometric estimates could serve to test whether CDPs have made an impact on inter-organizational networks, which are in turn held responsible for the effectiveness of the cluster program. This permits evaluators to isolate whether the impact on the performance indicator is due to network-related effects or to other factors. One of the current challenges in CDP evaluation is that, given the co-located nature of cluster firms, treated firms are highly likely to generate externalities that will also benefit non-treated firms. As Maffioli *et al.* (2011) put it, "it can be considered… that firms that share the same location have linkages. In this case, non-treated firms that are located in the same municipality of treated firms can be considered indirect beneficiaries." (p. 12). This complicates the evaluator's task because it makes the search for counterfactuals or control groups harder, given that non-treated firms will be indirect beneficiaries of the policy anyway. Nevertheless, we know from earlier research that co-located firms need not be linked to each other, and cluster research shows that spillovers are not evenly distributed in space (Giuliani, 2007). SNA helps to detect which firms have connections to the treated firms—and the types of connections—and hence can be indirect beneficiaries of the CDP. SNA also helps to identify local firms with no connections altogether. This is certainly critical information for policy evaluators, as it helps to open up the black box of the relationship between the CDP and its impact (Matta, 2010).

Box 1: The Importance of SNA for CDP Evaluation: Practical Examples

The theory and methodology underpinning SNA can stimulate policymakers' thinking—not only about CDP evaluation, but about CDP design and implementation. An advanced understanding of the advantages and limitations of different network positions and structures (Section 2.2) helps policymakers move beyond a generic network-stimulating policy perspective. Given network positions or structures can become the expected outputs of a project and be associated with given expected impacts. This would allow for the development of more narrowly defined cluster programs, with the caveat that network structures should not be "imposed" on the treated firms. In this respect, it is important to clarify here that **there is no single "optimal" structure or network position**. As discussed in Section 2, each type of position or structure conveys advantages and disadvantages. Hence, the efficacy or rewards of a specific network structure or position is likely to vary from context to context and cannot be identified in absolute terms. It is left to the evaluators and designers of cluster policies to judge whether targeting a specific structure (or a range of structures) brings some advantages to the whole policy development and evaluation process.

To stimulate thinking around this issue, we offer some stylized examples about the connection between network characteristics and CDP impact.

1) **Poverty eradication** and **even development.** If the expected impact is promoting the empowerment of poor people and increasing the inclusiveness of poor artisans in rural communities in the economy and society, then CDPs may have to focus on creating local **cliquish** networks, which promote trust and social capital and sanction opportunistic behaviors. Density, as well as cliquish structures, may be relevant network indicators to look at for output evaluation. As for impact evaluation, evaluators should analyze whether higher participation in local networking processes reduces poverty at the individual level.
2) **Well-organized value chain.** Policymakers may be interested in the promotion of a well-organized local value chain, coordinated by one or a few leading firms, which are expected to orchestrate the local value chain and connect it to global markets. Here, the desired output may be a **hierarchical network**, and CDPs may be interested in **increasing the centrality** of some of the leading firms, while leaving other firms (e.g., suppliers) in less central positions. If a hierarchical network is the desired output, selective connections should be promoted, while redundant linkages should be avoided. It may be the case of a supplier that does not need to be connected with many firms but simply with a few key clients. Next, evaluators should assess whether different types of centralities affect firm-level performance.
3) **Innovation**. CDPs may be oriented at promoting local innovation processes. **Local** cliquishness may favor the sharing of high-quality knowledge, thus generating opportunities for incremental innovation. At the same time, too many connections are considered detrimental for innovation, hence CDPs may be interested in encouraging connections only among partners that have something valuable to share. In this case, avoiding redundant ties is also an important aspect of CDPs. In this respect, in certain cases or contexts, connections with distant actors can be desirable (**small world**) as can promoting diversity in the nature of partners (**structural holes** and **brokerage**).

Section 3: SNA in Practice

3.1 Relational Data

SNA is generally considered to be something more than a mere methodological tool. It is first and foremost a distinct way to conceive the society and the economy. The network perspective differs in fundamental ways from standard social or behavioral science research and methods: "rather than focusing on attributes of autonomous individual units, the associations among the attributes, or the usefulness of one or more attributes for predicting the level of another attribute, the social network perspective views characteristics of the social units as arising out of the structural or relational processes or focuses on properties of the relational systems themselves" (Wasserman and Faust, 1994). Hence, the collection of *relational data* is critical for the analysis of social networks. In the context of industrial clusters, relational data can be of a different nature, depending on the content and characteristics of the CDP: it can either involve formal relationships, like the trade of goods or research-and-development (R&D) contractual agreements, or informal linkages, such as the transfer of tacit knowledge or any other type of informal collaboration. Also, in the collection of data, different types of actors may be targeted: only business firms or entrepreneurs in some cases, or also other actors, such as universities, non-profit organizations, government agencies, and international actors. Beside the collection of relational data, it is important to collect additional information about actors' attributes.

3.1.1 The Collection of Data

Relational data are collected by asking actors about their relationships with other actors, which they have to identify and name. It differs from other approaches, as it deliberately asks about relationships between identifiable actors, and not between the respondent and general categories or groups of actors—suppliers, clients, universities, etc. This clearly makes confidentiality agreements with the respondents of critical importance, as some interviewees may be unwilling to provide relational information that involves other actors. However, as the exponential growth of network research shows, obtaining network data is not an insurmountable problem (for a discussion on the limitations and caveats of this approach, see Section 3.3). In a restricted number of cases, the problem of confidentiality can be solved by relying on secondary data, such as co-patents, co-publication data, or data about formal participation in joint projects (Russo and Rossi, 2009; on a related issue, see also Ubfal and Maffioli, 2010). However, these types of data are neither always available nor meaningful in many clusters, as they capture only a fraction of the linkages existing between local actors. They are certainly to be exploited in places characterized by advanced R&D and scientific research activities, where formal ties can account for much of the connectivity of cluster firms and other actors. In the context of LAC countries' clusters, especially in low-technology industries, things may be different and relational data may need to be collected through direct interviews or observations. Collecting relational primary data in industrial clusters has generally taken place through direct interviews based on structured questionnaires, either in the context of CDP evaluation studies (Matta and Donadi, 2007; Matta, 2010) or, more generally, in that of academic research (e.g., Giuliani and Bell, 2005; Giuliani, 2007; McDermott *et al.*, 2009).

Depending on the size of the cluster population, it is possible to opt for different data collection methods. If the population size is limited (e.g., less than a hundred actors), the best option is to interview all the cluster actors by using a "roster-recall" method. This means that the questionnaire should be designed to include a complete list (roster) of the actors in the cluster and should ask about the existence, or the importance, of a given type of relationship the interviewed actor (called focal actor) has with all the rest in the roster. The questions may be formulated to collect either dichotomous relational data (existence/nonexistence of a tie) or valued data, which weight relationships on the basis of their importance, frequency, size, value, etc. Also, relationships can be mutual by design (e.g., a question about the geographical distance between two actors) or unidirectional (e.g., relationships where an asset is transferred from actor *i* to actor *j*). Box 2.1 reports some examples of questions based on the roster-recall method.

The collection of these types of relational data can be undertaken via face-to-face interviews or even Web questionnaires, depending on the nature and characteristics of the interviewees: for rural farmers, face-to-face interviews are the appropriate tool, while information and communications technology (ICT) managers may be happy and capable of answering through a Web survey. The roster-recall method applied to the whole population of cluster firms is the best and most recommended way to collect network data. It minimizes the risk of data loss due to a respondent's poor memory, as each interviewee has the complete list of other actors in the cluster to consult before answering questions about relationships. Also, this method permits the study of the full network, which is an important technical aspect, as it allows the analysis of actors' network positions (Section 2.2.1) and of the network's structural properties (Section 2.2.2).

Unfortunately, it is not possible in all cases to access the overall population of actors in a cluster, as the number of firms and other actors may be so high (e.g., several hundreds or thousands) that it is practically unfeasible to collect data on the whole population and use a roster-recall method. In such cases, different strategies can be adopted. One strategy is to identify criteria for sampling the network population. A random sampling is not recommended (Wasserman and Faust, 1994). Rather, select the relevant population on the basis of some justifiable criteria. For instance, perhaps only the largest firms will be considered in the analysis or, by contrast, only the smallest ones.

Alternatively, it is possible to select only firms occupying specific stages of the value chain (e.g., only assembly firms) or firms in selected industries only (in the case of several industries being present in the same cluster). The criteria for selecting the subpopulation are very important, as they influence the interpretation of results, and they should be chosen based on the characteristics of the local cluster as well as on the CDP evaluation objectives. Once the subgroup has been selected, it is possible to collect intra-subgroup relational data through the roster-recall method, following the approach mentioned earlier but including in the roster only the population of actors within the subgroup. In addition, it is possible to ask about relationships with the rest of the cluster population. As these may be too numerous to be included in a roster, or even not totally known prior to data gathering, the best way to collect these relationships is through a "free-recall" method. The free-recall method leaves the respondent free to name cluster actors without there being a predetermined list. In this case, questions can be formulated following either a "free choice" or a "fixed choice" design. In the latter case, respondents are given a fixed maximum number of ties to other actors to choose, so that they will name only the most important actors (e.g., most important five), while not providing relational information about other less important linkages (see Box 2.2 for examples). In the free-choice design, respondents are not given any such constraints on how many nominations to make, which has the advantage of including all the ties in spite of their relevance, but has the risk of inaccurate replies (see Box 2.3).

Collecting relational data based on a subset of the population is the only viable collection method in highly populated clusters. An alternative strategy is to sample some local actors and reconstruct the structure of the network around them. This is the so-called "ego-centered" network, which is defined as a partial network that is anchored around a particular node (individual or organization, called the "ego") and its direct contacts (alters). In some cases, it is possible to collect information about alters' linkages directly from the ego (Box 2.4). However, the best way to collect good and reliable "ego-centered" network data is to interview the ego first, and then use a snowballing approach to interview egos' alters about their relationships. Generally, this approach does not permit evaluators to map the full network, but only the networks surrounding the sampled actors. In some cases, it is also possible to adopt a snowballing approach and interview alters' alters, and to reconstruct through a bottom-up approach the relevant population, but this, again, is recommended for small populations only.

Box 2: Examples of Relational Questions

2.1 Roster-Recall
Advice seeking/giving
If you are in a critical situation and need technical advice, to which of the local firms mentioned in the roster do you turn? [Please rate the importance you attach to the knowledge linkage established with each of the firms according to its persistence and quality, based on the following scale: 0 = none; 1 = low; 2 = medium; 3 = high.]

Which of the firms in the roster do you think have benefited from technical support from this firm?
[Please rate the importance you attach to the knowledge linkage established with each of the firms according to its persistence and quality, based on the following scale: 0 = none; 1 = low; 2 = medium; 3 = high.]

Collaboration in marketing
Indicate the firms in the roster with whom this firm has collaborated for the development of a joint marketing initiative in the past two years. [Please rate the importance you attach to the collaboration according to the impact it had on your business activities, based on the following scale: 0 = none; 1 = low; 2 = medium; 3 = high.]

Collaboration as a result of the CDP
Indicate the firms in the roster with whom this firm has collaborated in the development of a new product for the domestic market as a direct consequence of the policy initiative [specify which]. [Please rate the frequency of collaborations according to the following scale: 0 = none; 1 = few times a year; 2 = monthly; 3 = weekly]

2.2 Free-Recall (Fixed Choice)
Contribution to product upgrading
Please name up to five local farmers whom you consider to have contributed to upgrading the quality of your crops.

University-industry linkages
Please indicate the names of up to 10 university researchers with whom you have interacted through at least one of the different activities listed below, in the past five years:

(i) joint research agreements (research undertaken by both parties);
(ii) consultancy work (commissioned by industry, not involving original research);
(iii) informal contacts (technical advice not based on a market transaction);
(iv) attendance at conferences with industry and university participation;
(v) other (specify______________________)

Innovative ties as a consequence of CDP
Please name up to five public research institutions with which this firm has established new innovative ties as a consequence of the CDP [normally a precise definition of "innovative tie" is provided].

2.3 Free-Recall (Free Choice)
Whom you trust (free choice)
Please name the entrepreneurs operating in the local cluster that you trust most [normally a precise definition of "trust" is provided].

Whom you trust (name generators)
Please name the local entrepreneurs to whom you would lend machinery for free.
Please name the local entrepreneurs to whom you would lend money.
Please name the local entrepreneurs that you would inform first about a new market opportunity.
(etc.)

Whom you trust more after the CDP (free choice)
Please name the entrepreneurs operating in the local cluster that you trust more as a consequence of the cluster policy initiative [specify which] + [normally a precise definition of "trust" is provided].

2.4 Ego-Centered Network
Question to the "ego"
Please name the firms in the cluster with whom this firm has collaborated for the solution of local environmental problems in the past three years.

Question to the "alters" (only about ego's alters—included in a roster)
Indicate the firms in the roster with whom this firm has collaborated for the solution of local environmental problems in the past three years.

3.1.2 Relational Datasets

To be analyzed, network data need to be organized in relational datasets. For small networks, relationships are expressed in *n* x *n* matrix form, with *n* being the number of actors in the network (sociomatrix). The rows and columns of the sociomatrix index the individual actors arranged in identical order. Each cell in the matrix reports the existence of a relationship between actor *i* and actor *j*. Depending on the question, relationships can be dichotomous (0, 1) or valued. If the relationship is based on mutual linkages, then the tie is considered "undirected" and the sociomatrix is symmetrical (e.g., the geographical distance between two actors or whether the two actors have a joint research agreement in common). In contrast, if the relationship is unidirectional, then the relationship is "directed" and the sociomatrix may be asymmetrical (e.g., in the case of advice networks, or of linkages formed for transferring resources and other assets). In Figure 3, we report the case of a valued and asymmetric sociomatrix. A different sociomatrix corresponds to each type of relationship. In the case of large and sparse networks (many actors but relatively few linkages), it is possible to use other approaches for the input of network data, such as DL (data language) files.[9]

Figure 3. Example of a Sociomatrix

	n1	*n2*	*n3*	*n4*	*n5*
n1	0	3	0	2	1
n2	5	0	2	3	5
n3	2	4	0	2	2
n4	3	3	3	0	3
n5	1	1	1	2	0

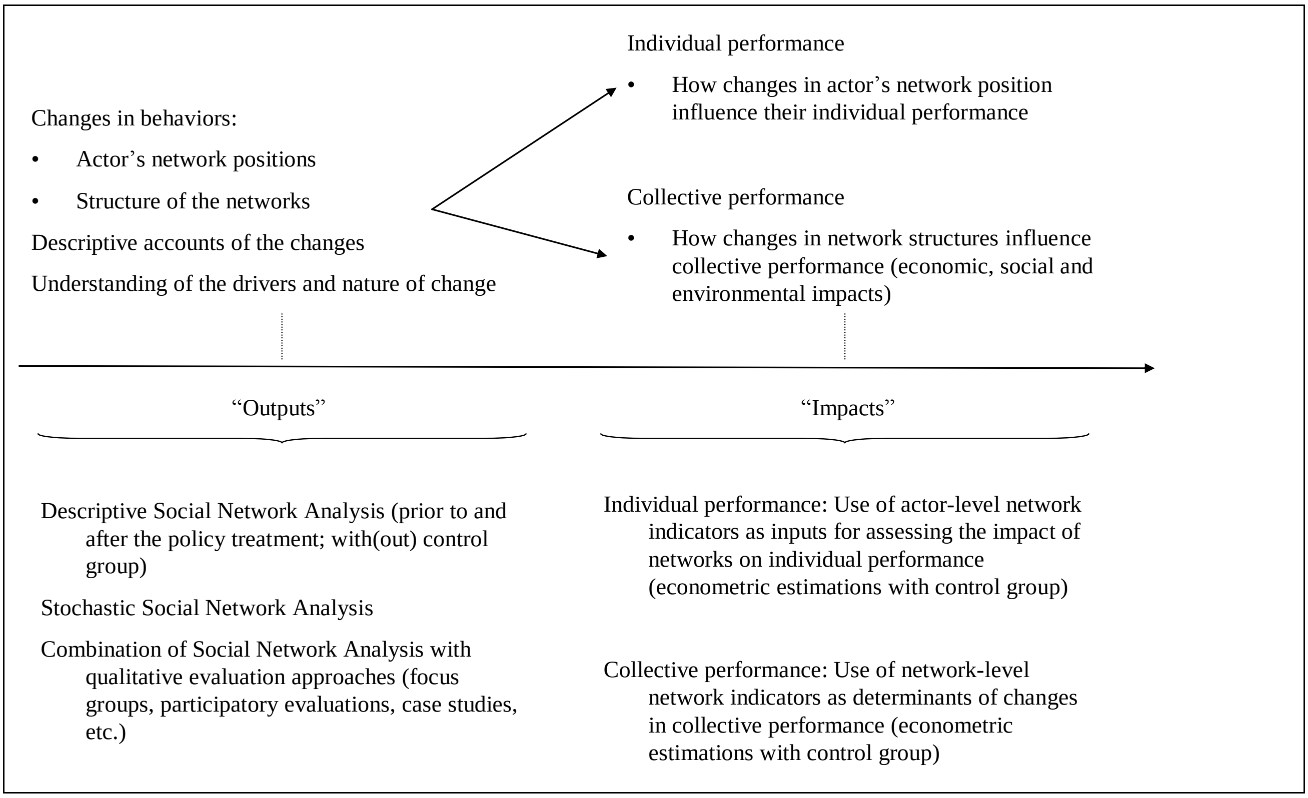
3.2 Analysis of Social Network Data for CDP Evaluation

As mentioned in Section 2.3, SNA is a versatile methodological tool. We envisage here two main SNA applications for CDP evaluation purposes (for a summary, see Figure 4). First, it can be used as a complement to more qualitative policy output evaluation approaches, which aim to *understand* rather than *test* the process underlying the policy, and evaluate whether the policy works smoothly and induces relevant changes to the targeted networking activities. This is essentially a first

[9] For instance, popular software for analyzing social networks, especially small ones, is UCINET, which allows relational data to be included as DL files, where only existing connections are included in txt file (see Borgatti *et al.*, 2002 and http://www.faculty.ucr.edu/~hanneman/nettext/C6_Working_with_data.html).

exploratory step in the evaluation, but it is nonetheless important to learn about policy processes and to interpret subsequent impact evaluations. In Sections 3.2.1 and 3.2.2, we provide a selective overview of measures and methods that are available to support output assessments of this type. Second, SNA indicators may be used as input in the evaluation of the impact of CDPs, which can be undertaken at the level of the individual organizations, and at the level of the whole community of firms and organizations populating a cluster, as discussed in depth in Section 3.2.3.

Figure 4 CDP Evaluation Roadmap Using SNA



3.2.1 Descriptive Network Analysis for Output Evaluation

It is certainly intriguing to know how a policy has changed the behavior or the choices made by treated actors with respect to whom they get connected to. Several questions are likely to occur to many CDP evaluators. One is whether networks change at all during and after the policy treatment, and whether the changes are different from those occurring in non-treated actors. A second one is whether these networks endure over time and are self-sustaining (Raines, 2002). A third possible question is whether changes in the network also come with changes in actors' positions: has a new central, powerful, and influential actor emerged (or disappeared) after the policy treatment? Another question could be whether the network that emerges at the end of the policy treatment has features that are coherent with the expectations and the design of the policy itself. More sophisticated evaluators may even want to know what has driven the changes in the network; that is, what micro-level strategies have been followed by the policy-targeted actors in the choice of their new ties, which have, in turn, lead to the emergence of a new network configuration. All these questions can be answered through descriptive and stochastic SNA.

Before turning more specifically to the measurements, we should clarify that there are different ways to approach these (and similar) questions. First, it is recommended that network data be collected prior to, during, and/or after the policy treatment. Although analyzing networks only at one point in time can provide useful information for evaluation, the best approach is to collect baseline data *prior* to the treatment and then *after* it, possibly even some years after the completion of the policy, to assess the self-sustainability of the network over time. Second, it is acceptable to look only at the treated cluster and network, without recurring to a control group, in which case qualitative interviews and focus groups should be used to provide an interpretative framework that can trace the logical connections between the changing network and the policy itself. However, collecting data about a control group allows for a more rigorous analysis on CDP impact, as discussed in Section 3.2.3. Third, through the relational questionnaires, and in parallel to questions about different types of relationships, it is advisable to ask interviewees for their perceptions about whether any of the newly formed relationships have been facilitated by the policy initiative (as also suggested in Box 2). Although this is likely to be a perceptive and highly subjective measure, it is relatively easy information to obtain once the questionnaire is answered, and it can give an appreciative idea about the direct role played by policies in shaping the network change.

Once evaluators are aware of these recommendations/caveats, a further critical analytical step is choosing the *right* mix of network indicators to be employed in the analysis. In Boxes 3 and 4, we include a selection of indicators that can be adopted to measure network position and network structure.[10] We do not recommend analysts be uncritical about their use. More often than not, SNA-based cluster (evaluation) studies tend to include long and tedious empirical lists of indicators—from different measures of centrality to different measures of network structure—but very little is actually done with them in terms of empirical analysis. Hence, after having had to digest a whole set of network measures and analyses, the reader often asks the inexorable "so what?" question. Based on the typical mistakes observed in studies using SNA in the analysis of clusters, our advice is to select those indicators that appear to make sense for informing the specific policy evaluation that is put in place. For instance, looking at structural holes (Box 3) may make sense only if there is a strong underlying motivation for this type of structural position to bring some benefits for the actors (e.g., in terms of creativity), or if such a position is one of the expected and desired outputs of the policy. Other centrality indicators may be more appropriate if the objective is to look at the change of power of some network actors. Likewise, density (Box 4) is always a good general indicator of connectivity, but it tells us very little about how a network is structured, and density indicators are not comparable across networks of different sizes. Also, the degree to which a density value is considered high or low varies greatly across networks and contexts. A value of 0.30 (for example) may be extremely high in some cases, and extremely low in others. Also, very dense

[10] Neither of the two boxes can be considered exhaustive. They are just illustrative examples of what measures could be available to describe networks, and the reader should be informed that new measures are being developed every day. In addition, a much more comprehensive review of network indicators can be found in specialized texts (e.g., Wasserman and Faust, 1994) and journals (e.g., *Social Networks; Journal of Mathematical Sociology*).

networks (indicator approaching 1) need not be always advantageous for their members, as they may host redundant and counterproductive ties. In a similar vein, once a given structure is detected in the network—suppose the network has a core-periphery structure (Box 4)—it does not make sense to try to fit the network into other structures as this may confuse the reader and add nothing to the evaluation report.

Box 3. Measures of Actors' Network Positions

3.1 Degree Centrality

Degree centrality is defined as the number of links incident upon a node (i.e., the number of ties that a node has). If the network is undirected, this is measured as the sum of linkages of firm *i* with other *j* actors of the network. The indicator can be standardised by *n*, with *n* being the number of nodes in the network:

$$DC_i = \frac{\sum_i xij}{n-1}$$

If the network is directed, two separate measures of degree centrality can be calculated:

- Out-degree centrality: measures the extent to which a tie originates from an actor.
- In-degree centrality: measures the extent to which a tie is incident upon an actor.

The indicator is computed on two alternative bases:

- Dichotomous: reflects the presence/absence of such a linkage.
- Valued: analyzes the value given to each linkage.

3.2 Betweenness Centrality

Betweenness centrality measures the number of shortest paths (geodesics(◘)) between *i* and *k* that an actor *j* resides on.

g_{jk} is the number of geodesics between *j* and *k*, and $g_{jk}(n_i)$ is the number of geodesics actor *i* is on.

$$BC(n_i) = \sum_{j<k} g_{jk}(n_i) \Big/ g_{jk}$$

The indicator can be standardized by dividing it by the number of pairs of actors not including n_i

$$St\_BC(n_i) = BC(n_i)/[(g-1)(g-2)/2]$$

3.3 Bonacich Centrality

An actor's centrality is a function of the centrality of those connected to that actor. Given an adjacency matrix A(◘), the Bonacich centrality of actor *i* (BoC_i), is given by:

$$BoC_i = \sum A_{ij}(\alpha + \beta * BoC_j)$$

where α and β are parameters. The centrality of each actor is therefore determined by the centrality of the vertices it is connected to. The value of α is used to normalize the measure, and the value of β is an attenuation factor that gives the amount of dependence of each vertex's centrality on the centralities of the vertices it is adjacent to. The parameter β is selected by the user: negative values should be selected if an individual's power is increased by being connected to vertices with low power, and positive values should be selected if an individual's power is increased by being connected to vertices with high power.

3.4 Ego-Centered Measures

3.4.1 Size of Ego-Net

The size of the ego-net measures the number of actors that have direct connections with the ego. It is a measure of degree centrality applied to ego-centered networks.

3.4.2 Local Cliquishness

Local cliquishness is generally measured through the clustering coefficient. The clustering coefficient *CCi* for an actor *i* is the proportion of links between the vertices within its neighborhood (alters) divided by the number of possible links that could exist between them. This reveals the extent to which a firm's ego-centered network is formed by alters that are connected to each other. It ranges between 0 and 1, being 1 when alters are all connected to each other, and 0 when they are not.

$$CCi = \frac{\left|\{e_{jk}\}\right|}{k_i(k_i - 1)}$$

e_{jk} is the number of linkages existing between the actors in the neighborhood of actor *i*, while k_i is the number of actors in the neighborhood. For a directed graph, it is possible to calculate the in- and out-neighborhood measures.

3.4.3. Structural Holes

The measure of access to structural holes is calculated as 1 minus the firm's *constraint* score in cases where the constraint is non-zero, and 0 for all other cases, because a zero score arises only when an actor is unconnected to others, and thus has no access to structural holes.

Consistent with Burt (1992), the measure is:

$$Constr = x_{ji} + \sum x_{iq} * x_{jq}, \qquad q \neq i, j$$

where x_{ji} equals the number of direct ties from *j* to *i* and $\sum x_{iq} * x_{jq}$ is the sum of indirect ties from *j* to *i* via all *q*.
Note: All terms identified with the (◘) symbol are described in the Appendix.

Box 4. Measures for Detecting Network Structures

4.1 Density

Network density (*ND*) is defined as the proportion of possible linkages that are actually present in a graph. *ND* is calculated as the ratio of the number of linkages present, *L*, to its theoretical maximum, *n(n-1)/2*, with *n* being the number of nodes in the network (Wasserman and Faust, 1994):

$$ND = \frac{L}{[n(n-1)/2]}$$

4.2 Cohesive Subgroups

4.2.1 Cliques, *n*-cliques, *n*-clans

A clique is a maximal complete(◘) subgraph of three or more nodes, all of which are adjacent(◘) to each other. A *n*-clique is a maximal subgraph in which the largest geodesic distance (◘) between any two nodes is no greater than *n*.

Hence, when *n* = 1, the subgraph is a clique, when *n* = 2 the subgraph is composed of members that need not be adjacent, but need to be reached through at most one intermediary. The *n*-clique is a "relaxed" version of a clique, given that large cliques are not easy to find in real networks. In some cases, *n*-cliques may not be very cohesive, especially for high numbers of *n*. Hence, other subgroup formulations are possible, which are in-between cliques and *n*-cliques, such as *n*-clans, which are *n*-cliques whose diameters are no greater than or equal to *n* (for more subgraph specifications, see Wasserman and Faust, 1994).

4.2.2 *k*-cores

A *k*-core is a type of subgraph defined on the basis of the actors' degree centrality. A *k*-core is a subgraph in which each node is adjacent to at least a minimum number, *k*, of the other nodes in the subgraph. On this basis, for instance, a 4-core is a subset of actors that have a degree centrality of at least 4.

4.2.3 Small Worlds

Small worlds are generally described as networks where the clustering coefficient is high relative to its random limit, yet the average shortest path length is as "small" as possible.

4.2.4 Core-Periphery

Core-periphery analysis allows the identification of a cohesive subgroup of core firms and a set of peripheral nodes that are connected with the core, but that do not connect to other peripheral nodes (Borgatti and Everett, 1999). A simple measure of how well the real structure approximates the ideal is given by the following two equations:

$$\rho = \sum_{i,j} a_{ij}\delta_{ij}$$

$$\delta_{ij} = \begin{Bmatrix} 1, if \\ c_j = CORE; \\ or \\ c_j = CORE \\ 0, otherwise \end{Bmatrix}$$

where a indicates the presence or absence of a tie in the observed data, c_j refers to the class (core or periphery) actor j is assigned to, and δ_{ij} indicates the presence or absence of a tie in the pure core-periphery structure. For a fixed distribution of values, the measure achieves its maximum value when and only when A (the matrix of a) and Δ (the matrix of δ_{ij}) are identical, which occurs when A has a perfect core-periphery structure.

4.3 Scale-Free

A scale-free network is a network whose degree distribution follows a power law, at least asymptotically. That is, the fraction P(n) of nodes in the network having n connections to other nodes goes for large values of n as

$$P(n) \approx cn^{-\gamma}$$

where c is a normalization constant and γ is a parameter whose value is typically in the range $2 < \gamma < 3$, although occasionally it may lie outside these bounds.

Note: All terms identified with the (◘) symbol are described in the Appendix.

Once the key and relevant indicators have been selected, evaluation may proceed by comparing them over time (prior to, during and/or after the policy treatment) to measure whether any significant change has occurred. Changes can occur in the structural positions of actors, and policy evaluators may be interested in monitoring the behavior of some specific actors (e.g., a subsidiary of a multinational company, a research laboratory, a leading firm). By comparing degree-centrality indicators over time, we can appreciate whether some actors have become more (or less) central relative to others, or whether they have acquired some bridging properties between other actors in the network that were previously unconnected. If, for example, the policy objective was to integrate academic research with the innovation activities of cluster firms, then looking at how the centrality of relevant university departments has changed may provide insights about the formation of new university-industry ties, accounting not only for the direct ties, but also for the new ties formed by the firms that have direct connections with the university (for an application, see Giuliani and Arza, 2009). Also, because ties can be valued by the respondents according to their relevance or frequency, it is possible to explore whether ties have become stronger over time. The options are numerous and depend very much on what is to be evaluated.

Furthermore, by comparing the evolution of networks of the same set of actors over time, we learn about the changes in the collective behavior of cluster firms. It is possible to know whether a given network has become more (or less) egalitarian, efficient, redundant, fragmented, etc., and therefore, whether some of the policy goals have been achieved at the collective level. As suggested in Box 5, SNA also allows us to explore the correlation between different types of networks (based on the same set of actors), and run regressions with network data as dependent and independent variables. This application can be useful, for appreciative purposes, if the evaluation project collects data about relations that have been explicitly and directly generated by the CDP initiative (we call this a "policy network"). An example is that of policies that involve the creation of a formal consortium of firms. Using quadratic assignment procedure (QAP) and multiple regression QAP (MRQAP) techniques, we can assess whether there is a correlation between the policy network and any other more spontaneous network that may have been formed *after* the formation of the policy network, controlling, for instance, for pre-existing networks and other factors such as the geographical distance between firms. As mentioned earlier, this type of analysis does not guarantee that the formation of new linkages is necessarily *caused* by the policy network.

Box 5. Correlations and Regressions of Networks (QAP and MRQAP)

To analyze the correlation among different networks, scholars normally refer to QAP (Krackhardt, 1987). This methodology is based on a two-step algorithm. First, it computes the correlation between corresponding cells of the two data matrices. Second, it randomly permutes the rows and columns of one matrix and recomputes the correlation. The second step is carried out 2,500 times to compute the proportion of times that a random measure is larger or equal to the measure calculated in the first place. A proportion lower than 0.05 (*p*-value) suggests that the relationship between the matrices is unlikely to have occurred by chance (Borgatti, Everett, and Freeman, 2002). Different types of correlation indicators can be computed. Two of the most frequently used are the Pearson coefficient and the Jaccard coefficient. The former is the conventional Pearson coefficient, but applied to network data; the latter is a measure of the proportion of matches when at least one of the observations has a 1, which means that it measures the co-occurrence of linkages for two different matrices considering only the existing linkages. The QAP correlation is carried out considering pairs of networks.

On the same principles lies MRQAP, which estimates the joint effect of more than one network on a dependent network, such that:

$$Y = \alpha + \beta 1X1 + \beta 2X2 + \beta 3X3 + \varepsilon$$

Y is the dependent sociomatrix, while X1, X2, X3 are other networks that are considered to influence the behavior of Y. α is the constant term and ε is the residual matrix. See Dekker, Krackhardt, and Snijders (2005) for further details on this regression approach.

3.2.2 Stochastic and Dynamic Network Analysis for Output Evaluation

So far, we have focused on deterministic analysis, which aims to describe the structure of a network as it is observed at any point in time, and to compare networks and network indicators over time. However, in some cases, it may be appropriate to adopt a statistical approach to network analysis to test, in probabilistic terms, whether an observed structure fits with an expected theoretical network structure, and to verify that the observed behavior is not given by chance. The importance of statistical testing, with respect to pure descriptive analysis, is based on two main considerations (Robins *et al.*, 2007). First, networks are the representation of complex decision processes, and stochastic models capture both the regularities in the processes that give rise to new linkages, and the possibility that part of a network structure is due to chance. Second, stochastic models identify the micro-level processes that lead to the formation of a given structure. In particular, they permit evaluators to discern between processes that are socially derived (endogenous to the network) from those that are related to actors' internal characteristics (exogenous to the network). Examples of endogenous network effects are the search for reciprocated ties, or the tendency of actors to form closed networks (see Table 4.1 for further examples). Exogenous network effects may explain how the heterogeneity of actors influences the formation of ties, for example, based on the similarity of actors (see Table 4.2 for further examples).

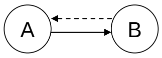
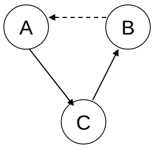
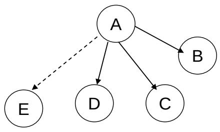
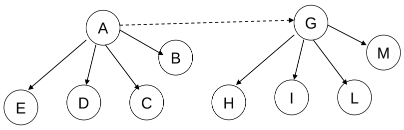
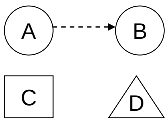
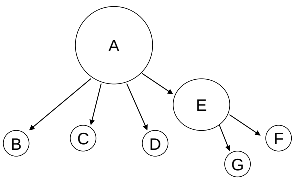
Furthermore, sophisticated evaluators may have an interest in obtaining robust estimates of the micro-level effects that have produced a given network structure. When networks are available at one point in time (e.g., at the end of the policy treatment), this type of analysis can be undertaken through exponential random graph models, also known as *p**models (Frank and Strauss, 1986; Robins, Pattison, and Wasserman, 1999; Snijders *et al.*, 2006; Robins *et al.*, 2007).[11] In practice, this means inferring whether a given effect (as in Table 4) is present in the observed graph to a greater or lesser extent than expected by chance, controlling for other effects. To test this, random networks are generated through Monte Carlo simulation methods from a starting set of parameter values and compared with the observed network. If panel network data are available (prior to and after the policy treatment), then stochastic actor-oriented models for network dynamics developed by Snijders and colleagues (Snijders, 2001; 2005; Snijders *et al.*, 2010) can be applied. Models for network dynamics look at whether the change of a network from time t to time $t + 1$ is the result of the simultaneous micro-level endogenous and exogenous network effects mentioned earlier (Table 4).

Through this methodology, evaluators who have access to panel network data can dig into the micro-level motivations and strategies that have led actors in the network to form new ties and to give rise to changes in the network structure. The usefulness of this methodology for CDP evaluation depends on the extent to which evaluators want to understand the *process* of network change and back up this understanding through statistical analysis. For example, this analysis can tell the extent to which the search for reciprocated ties matters, relative to the similarity of actors' characteristics, on the formation of new linkages. In a recent study, Giuliani (2010) finds that the evolution of the knowledge network in a Chilean wine cluster is due to two concurrent micro-level effects: the cohesive effect (reciprocity and transitivity) and the capability effect, as she finds that firms with weak capabilities are persistently undersocialized. If network data about policy-induced linkages are available (e.g., collaborations for joint product development funded by the program), it is also possible to test statistically whether networks formed as a direct output of the program also affect other types of networks, which may have been formed later on by cluster members.

---

[11] The mathematical treatment of these types of models goes beyond the scope of this paper.

Table 4 Endogenous and Exogenous Network Effects in Stochastic Network Models

	Illustration	Explanation
1. Endogenous mechanisms:		
Reciprocity		The formation of new ties is based on the search for reciprocation.
Transitive triplets		A new tie is more likely to occur between A and B if A and B are tied to a common actor (C)
Preferential attachment		This is tested through the out-degree popularity effect. A positive and significant effect reflects the tendencies for actors with high out-degrees (i.e., outgoing ties) to attract extra incoming ties.
Assortativity		This effect tests for the existence of actor preference to form new ties with actors with similar ingoing or outgoing ties.
2. Exogenous mechanisms:		
Similarity		A new tie occurs more often between firms with similar values in a given attribute.
Activity		A new tie is more likely to occur when firms have stronger (weaker) values on a given attribute.
Note: The mechanisms included in this table are only for illustrative purposes. Additional mechanisms exist and can be tested.		

3.2.3 SNA as a Fundamental Input for Impact Evaluation

Descriptive and stochastic SNA cannot intrinsically lead to a full-fledged evaluation of CDP impact (Schmiedeberg, 2010). However, as mentioned in Section 2.3, the real breakthrough in impact evaluation comes from the combination of SNA with other quantitative methods for policy impact evaluation, such as quasi-experimental approaches with constructed controls (Oldsman and Hallbert, 2002; Adam, 2006). In particular, two types of opportunities are present: 1) the individual impact assessment, whereby the impact of network changes on actor-level performance is examined; and 2) the collective impact assessment, which instead looks at the impact that CDPs have on the community of firms (and other organizations/actors) in a cluster (for an illustration, see Figure 4).

To undertake individual impact assessment, indicators of actor-level network position need to be included as independent variables in impact assessment econometric estimations with (quasi-) experimental design, and indicators of firm-level performance need to be included as dependent variables (Maffioli *et al.*, 2011). With this approach, evaluators may test whether an improvement in performance is due to the way an actor is connected to other local actors or to other types of effects. Hence, rather than taking for granted the presence of a network effect deriving from the CDP, we would test this effect explicitly. This constitutes advancement in the evaluation of cluster policies, which generally fail to assess directly the connection between changes in business behavior related to connectivity and networks and performance (Raines, 2002). Moreover, it is possible, in principle, to have a fine-grained look at what types of network positions (Section 2.2) are most likely associated with performance. Firms are expected to react differently to a given CDP, and thereby make different connectivity choices resulting in a different positioning within the network. For instance, some actors may form dense and cliquish networks around them, while others will bridge structural holes. Through the above-mentioned econometric approaches, it becomes possible to test whether the enhanced performance is due to one type of position or the other, and (through interaction effects) to test whether one type of position, *combined with* certain characteristics of the firm, is most likely to generate an improvement in the performance.

Most importantly, this analysis would inform policymakers about *what type* of network position is best associated with performance: it would make visible what is normally invisible. In addition, this type of measure will facilitate the analysis of the direct impact of CDPs on the treated actors, through the indirect impact on: 1) non-treated actors that have connections with the treated actors; and 2) non-treated actors with no connections to the beneficiaries. This, as discussed at length in Maffioli *et al.* (2011), will allow for a much finer estimation of CDP impact.

The second way in which SNA can be used for impact evaluation is at a more aggregate level. Rather than focusing on individual actors (a firm, a private or public organization, etc.) as the unit of analysis, the focus can be on the whole cluster or region. A key CDP objective is often to improve the performance of a *whole* community of entrepreneurs and firms, not simply that of individual actors.[12] In this case, the unit of analysis is the cluster or the region, and different cluster-level performance measures may be explained by measures of network structure. The feasibility of this approach is conditional on the existence of a sufficient number of observations, which means that extensive and comparable data collection needs to be undertaken in different cluster contexts within the same country or within a group of countries. Having these types of data requires a long-

[12] As social and environmental concerns have become progressively more central in the policy-making discourse, performance at the community level need not be measured only via economic indicators, but may also include measures of social and environmental impact.

term investment by policy-making agencies, which need to standardize the method of data collection and analysis and plan it so as to avoid too much inter-cluster heterogeneity in the way data are collected and analyzed. The use of cluster-level data will permit evaluators to test whether, for instance, artisan clusters really need to increase the overall density of the network to reduce poverty. Similarly, evaluators could test whether a very dense network has a positive impact only in clusters with given characteristics, such as size, distance from frontier technologies, distance from main markets, etc.

This type of analysis offers an extraordinary learning opportunity for policymakers, as understanding what type of network structure is best for achieving the objectives of the policy can orient subsequent policy-design processes, and foster the development of a desired network structure. In fact, different types of policy measures are needed to achieve different types of network structures. For instance, to increase the density of local linkages, the United Nations Industrial Development Organization (UNIDO) makes use of a local broker, whose aim is that of making trustful connections among all actors in the local cluster, thus pursuing maximum connectivity (Pietrobelli, 2009). But there are contexts where high density is not what policies should look for. We have seen that selective and more efficient networks promote radical innovations (Section 2.2). On this basis, some CDPs may want to strengthen the linkages between key actors in the local cluster, particularly between those that have higher innovative potentials. Hence, in this case, competitive bidding schemes for joint R&D collaborations may be a more valuable policy tool. Understanding what type of network structure is best suited to achieve the objectives of the policy is of fundamental importance for informing policy design.

To conclude, SNA indicators, both at the firm and at the cluster level, can be applied in impact evaluation exercises. This requires evaluators to follow the standard methodology for impact assessment. Among other things, this means that the collection of data, including relational data, from a control group is needed. However, so far, there has been very little work in this direction (Maffioli, 2005; Ubfal and Maffioli, 2010; although their focus is not on cluster policies), and the few CDP evaluation studies that exist use SNA as a descriptive tool only. The potential of combining SNA with econometric approaches to impact evaluation is huge.

3.3 Caveats in the Application of SNA

One of the objectives of this paper is to give a full overview of the applicability of SNA to CDP evaluation processes. However, to be able to fully gauge the benefits of this methodological approach, the reader should be informed about its limitations and caveats, which are discussed in the remainder of this section.

First, *obtaining full network data may be troublesome*. In earlier sections, we have stressed that one of the advantages of SNA is that it allows the analysis of the network structure of a *whole* community of firms. However, in practice, it is not always possible to collect or access full network data, as this implies that the *whole population* of actors (or a selected subpopulation, see Section 3.1) within a cluster is interviewed and provides reliable information about their connections. The natural shortcut of ego-centered network data collection does not permit evaluators to map the full network, as only the ego's local neighborhood is available, and this information can be used only for individual impact assessment analysis, not for collective impact assessments. One of the problems with full network data is that non-respondents may severely distort data, as a network map may be very misleading if the central actor is not included (Borgatti and Molina, 2003). To avoid non-responses, we recommend that the design of a program includes SNA as an impact assessment methodology, so that the beneficiaries of the CDP are well informed about data requirements prior to the start of the program, and the necessary data are collected. Even in this case, however, problems remain regarding the construction of the control group. It is hard to guarantee a 100-percent response rate in the control group, which is composed of firms or other actors that have no obligation to and may have no interest in participating in the survey. If this is the case, impact analysis may be more difficult to realize unless ad hoc proxies (e.g., based on ego-centered networks) are identified in the control group, which can be compared with the observed network data of the treated cluster.

Second, *SNA imposes some ethical considerations*. Unlike conventional methodological approaches, anonymity at the data collection stage is not possible, as respondents have to report and name others with whom they have established relationships (Borgatti and Molina, 2003). Hence, a respondent may report on relationships with other actors, which may not wish to be named. This aspect is particularly problematic for sensitive relational data (e.g., "Who do you bribe in the government to obtain business licences?" "To whom do you transfer sensitive information about this firm?" "What strategic alliances do you plan to undertake in the next five years?"). The collection of data should bear these ethical issues in mind. First, confidentiality should be guaranteed by specifying that none of the relational information will be disclosed to other respondents and that network maps will not appear with the names of the actors (unless differently agreed with the respondents). Second, relational questions should be formulated to avoid as much as possible sensitive or highly strategic information, which the respondent may be unwilling to provide.

Section 4. Conclusions

This paper contributes to the relatively recent but growing literature on the impact evaluation of CDPs. In light of the enormous increase in these types of policies worldwide, policymakers and program managers are now in need of a well-developed methodological toolkit for their evaluation, but few methods and solid evaluations are available. In this paper, we consider one aspect of CDPs that is often at the core of these types of policies: the formation and/or strengthening of inter-organizational networks. We argue that prior evaluation studies have failed to measure network-related concepts appropriately, and we propose an alternative treatment of such concepts based on the application of SNA. The paper provides an overview of the advantages of SNA, and it is meant to be primer on how SNA can be used in CDP evaluation. In particular, it discusses how SNA can be applied: first, in combination with qualitative evaluation studies (e.g., Diez, 2001); second, in quantitative exercises of CDP impact evaluations (Maffioli *et al.*, 2011). We also highlight the caveats inherent to this methodological approach. To conclude, the paper intends to be a methodological term of reference for those wishing to apply SNA in the context of CDP evaluation. However, it is not a practical and ready-to-use toolkit for practitioners, and although it has tried to cover as much as possible the key SNA methodological methods and issues, readers interested in the application of SNA methods should also refer to specialized and comprehensive texts.

References

Adam S. 2006. Evaluating Social Fund Impact: A Toolkit for Task Teams and Social Fund Managers, *Social Protection Discussion Paper n. 0611, The World Bank*: Washington DC.

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

Aragón, C., MJ Aranguren, C. Iturrioz, *et al.* 2009. *A new step in cluster policy evaluation in the Basque country*. Orkestra Working Paper Series in Territorial Competitiveness Number 2009-01 (EN), Orkesta, Basque Institute of Competitiveness, Spain.

Baker, J. 2000. *Evaluating the Impact of Development Projects on Poverty. A Handbook for Practitioners*. Washington, D.C.: The International Bank for Reconstruction and Development/The World Bank.

Barabasi, A-L, R. Alber. 1999. Emergence of scaling in real networks. *Science* 286: 509–12.

Baruj G. 2007. Testeo en Campo de una Metodología Común de Monitoreo y Evaluación para

Baum, JAC, AV Shipilov, TJ Rowley. 2003. Where do small worlds come from? *Industrial and Corporate Change* 12(4): 697–725.

Bell, GG. 2005. Clusters, networks, and firm innovativeness. *Strategic Management Journal* 26: 287–295.

Bellandi, M., A. Caloffi. 2010. Architettura delle reti di innovatori finanziate dalla Regione Toscana (2000-2006). In *Processi di Innovazione e Sviluppo Locale. Teorie e Politiche*, eds. M. Russo and Donzelli. Rome.

Bonacich P. (1987) Power and centrality: a family of measures, *The American Journal of Sociology*, 92: 1170 - 82.

Borgatti, SP, MG Everett. 1999. Models of core/periphery structures. *Social Networks* 21, 375–395.

Borgatti, SP, MG Everett, LC Freeman. 2002. *Ucinet 6 for Windows*. Harvard: Analytic Technologies.

Borgatti, S. P., Everett, M.G. 2006. A graph-theoretic framework for classifying centrality measures. *Social Networks* 28(4): 466-484.

Borgatti, SP, JL Molina. 2003. Ethical and strategic issues in organizational social network analysis. *Journal of Applied Behavioural Science* 39(3): 337–349.

Borras, S., D. Tsagdis. 2008. *Cluster Policies in Europe: Firms, Institutions, and Governance*. Cheltenham, U.K., and Northampton, MA.: Edward Elgar.

Burt, RS. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, MA.: Harvard University Press.

-----. 2001. The social capital of structural holes. In *New Directions in Economic Sociology*, eds. MF Guillen, R. Collins, P. England, M. Meyer. New York: Russell Sage Foundation; 201–249.

Cattani, G., S. Ferriani. 2008. A Core/Periphery Perspective on Individual Creative Performance: Social Networks and Cinematic Achievements in the Hollywood Film Industry. *Organization Science* 19(6): 824–844.

Coleman, JS. 1988. Social capital in the creation of human capital. *American Journal of Sociology* 94: 95–120.

Condo, A., G. Monge. 2002. *Promoción de clusters en América Latina: la experiencia del CLACDS-INCAE*. Documentos en proceso del Centro Latinoamericano para la Competitividad y el Desarrollo Sostenible (CLACDS) del INCAE. Costa Rica.

Davis, GF, M. Yoo, WE Baker. 2003. The small world of the American corporate elite, 1982–2001. *Strategic Organization* 1:301–26.

Davis, R. 2003. *Network perspectives in the evaluation of development interventions: More than a metaphor*. Paper presented at EDAIS Conference, November 24–25, New Directions in Impact Assessment for Development: Methods and Practice.

De Solla Price, D. 1976. A general theory of bibliometric and other advantage processes. *Journal of American Society for Information Science* 27: 292–306.

Dekker, D., D. Krackhardt, TAB Snijders. 2005. *Sensitivity of MRQAP Tests to Collinearity and Autocorrelation Conditions*. Working Paper Series on Research in Relationship Management. Nijmegen School of Management, University of Nijmegen.

Diez, MA. 2001. The Evaluation of Regional Innovation and Cluster Policies: Towards a Participatory Approach. *European Planning Studies* 9(7): 907–923.

Europe INNOVA. 2008. *The concept of clusters and cluster policies and their role for competitiveness and innovation: Main statistical results and lessons learned*. Commission Staff Working Document SEC (2008) 2637. European Communities, Luxembourg.

----. 2002a. *Regional Clusters in Europe*. Observatory of European SMEs, N. 3. Enterprise Directorate-General, Brussels.

----. 2002b. *Final report of the expert group on Enterprise clusters and networks*. European Commission, Enterprise Directorate-General, Brussels.

Fernandez-Mateo, I. 2007. Who Pays the Price of Brokerage? Transferring Constraint through Price Setting in the Staffing sector. *American Sociological Review* 72, 2: 291–317.

Ferraro, C. 2010. *Clusters y políticas de articulación productiva en América Latina*. Comisión Económica para América Latina y el Caribe (CEPAL), Naciones Unidas: Santiago de Chile.

Fondo Multilateral de Inversiones (FOMIN). 2011. *Strategy for learning, communication and catalyzing. Expand the adoption of effective models and strategies for enabling MSME development at regional/local level*. Washington, D.C.: Inter-American Development Bank.

----. 2010. *Manual del sistema de monitoreo de la familia de proyectos de promoción de la competitividad local (SIMPLEd)*. Washington, D.C.: Inter-American Development Bank.

Frank, O., D. Strauss. 1986. Markov Graphs. *Journal of the American Statistical Association* 81: 832–42.

Freeman, C. 1991. Networks of innovators: A synthesis of research issues. *Research Policy* 20: 499–514.

Freeman, L. 1979. Centrality in Social Networks Conceptual Clarification. *Social Networks*, 1(1978/79): 215–239.

Fromhold-Eisebith, M., Eisebith G. 2005. How to institutionalize innovative clusters? Comparing explicit top-down and implicit bottom-up approaches, *Research Policy*, 34 (8): 1250-1268.

Fromhold-Eisebith, M., Eisebith G.. 2008. Looking Behind Facades: Evaluating Effects of (Automotive) Cluster Promotion. *Regional Studies*, 42: 10, 1343–1356.

Ganne, B., Y. Leclerc. 2009. *Asian Industrial Clusters, Global Competitiveness and New Policy Initiatives.* Hackensack, N.J., and Singapore: World Scientific.

Gargiulo, M., M. Benassi. 2000. Trapped in your own net? Network cohesion structural holes, and the adaptation of social capital. *Organization Science* 11(2) 183–196.

Giuliani, E. 2007. The selective nature of knowledge networks in clusters: evidence from the wine industry. *Journal of Economic Geography* 7: 139–168.

----. 2008. *What drives innovative output in emerging clusters? Evidence from the wine industry*. SPRU Electronic Working Paper Series 169, University of Sussex, SPRU – Science and Technology Policy Research.

----. 2010. *Network dynamics in regional clusters: The perspective of an emerging economy*. Industry Studies Association Working Papers Series 2010-09.

Giuliani, E., M. Bell. 2005. The micro-determinants of meso-level learning and innovation: evidence from a Chilean wine cluster. *Research Policy* 34(1): 47–68.

Giuliani, E., C. Pietrobelli, R. Rabellotti. 2005. Upgrading in global value chains: Lessons from Latin American clusters. *World Development* 33(4): 549–573.

Giuliani, E., V. Arza. 2009. What drives the formation of 'valuable' university-industry linkages? Insights from the wine industry. *Research Policy* 38: 906–921.

Giuliani, E., R. Rabellotti. 2011. Bridging researchers and the openness of wine innovation systems in Chile and South Africa. In *Innovation and Technological Catch-Up: The Changing*

Geography of Wine Production, eds. E. Giuliani, A. Morrison, R. Rabellotti. Cheltenham, U.K.: Edward Elgar.

Gould, RV. 2002. The origins of status hierarchies: A formal theory and empirical test. *American Journal of Sociology* 107(5): 1143–1178.

Gould, RV., RM Fernandez. 1989. Structures of Mediation: A Formal Approach to Brokerage in Transaction Networks. *Sociological Methodology* 19: 89–126.

Grabher, G. 1993. *The Embedded Firm. On the Socioeconomics of Industrial Networks*. London, New York: Routledge.

Granovetter, M. 1985. Economic action and social structure: the problem of embeddedness. *American Journal of Sociology* 91: 481–510.

Guillén, M.F. 2000. Business groups in emerging economies: A resource-based view. *Academy of Management Journal* 43(3): 362–380.

Gulati, R. 1995. Does familiarity breed trust? The implications of repeated ties for contractual choices in alliances. *Academy of Management Journal* 38(1): 85–112.

Hospers, GJ., P. Desrochers, F. Sautet. 2009. The next Silicon Valley? On the relationship between geographical clustering and public policy. *International Entrepreneurship Management Journal* 5:285–299.

Huggins, R. 2001. Inter-firm network policies and firm performance: evaluating the impact of initiatives in the United Kingdom. *Research Policy* 30Ž2001: 443–458.

Ibarra, H., SB Andrews. 1993. Power, social influence and sense making: Effects of network centrality and proximity on employee perceptions. *Administrative Science Quarterly* 38: 277–303.

Ketels, C., G. Lindqvist, O. Sölvell. 2006. *Cluster Initiatives in Developing and Transition Economies*. Stockholm School of Economics, Sweden: Center for Strategy and Competitiveness.

Khanna, T., JW Rivkin. 2001. Estimating the performance effects of business groups in emerging markets. *Strategic Management Journal* 22: 45–74.

Kogut, B., G. Walker. 2001. The small world of Germany and the durability of national networks. *American Sociological Review* 66: 317–35.

Krackhardt, D. 1987. QAP partialling as a test of spuriousness. *Social Networks* 9: 171–186.

Landabaso, M. 2000. EU policy on innovation and regional development. In *Knowledge, Innovation and Economic Growth: The Theory and Practice of Learning Region*, eds. F. Boekma *et al*. Cheltenham: Edward Elgar.

Laumann E.O., Pappi F.U. 1976. *Networks of collective action*. New York: Academic Press.

Laursen, K., A. Salter. 2006. Open for innovation: The role of openness in explaining innovation performance among U.K. manufacturing firms. *Strategic Management Journal* 27 131–150.

Lazega, E., L. Mounier, T. Snijders, *et al*. 2010. Norms, status and the dynamics of advice networks: A case study. *Social Networks*, forthcoming.

Lundquist, P., D. Power. 2002. Putting Porter into Practice? Practices of Regional Cluster Building: Evidence from Sweden. *European Planning Studies* 10(6): 685–704.

Luce, RD., AD Perry.1949. A method of matrix analysis of a group structure. *Psychometrika* 14: 95–116.

Maffioli, A., C. Rodriguez Mosquera, R. Stucchi. 2011. *A methodological approach to the impact evaluation of cluster development programs*. Mimeo.

Markusen, A. 1996. Sticky places in slippery space: a typology of industrial districts. *Economic Geography*, 72(3): 293–313.

Marshall, A. 1920. *Principles of Economics*. London: MacMillan.

Martin, R., P. Sunley. 2003. Deconstructing clusters: chaotic concept or policy panacea? *Journal of Economic Geography* 3: 5–35.

Maffioli, A. 2005. *The Formation of Network and Public Intervention: Theory and Evidence from the Chilean Experience*. Centre for research on Latin American Studies and Transition Economies, ISLA Working Paper 23. Bocconi University, Milan.

Matta, A. 2010, Reconstruyendo la trama associativa. In *Desarrollo de cadenas productivas, clusters y redes empresariales*, ed. F. Mitnik. Fondo Multilateral de Inversiones, Banco Interamericano de Desarrollo. Agencia de Desarrollo de la Ciudad de Córdoba (Argentina). In press.

Matta, A., L. Donadi. 2007. *Desarrollo de un modelo experimental para el monitoreo y la evaluacion de la cooperacion entre empresas en el programa de desarrollo de cadenas productivas. El caso Cadena informatica y electronica*. Mimeo.

McDermott, G., R. Corredoira, G. Kruse. 2009. Public-private institutions as catalysts of upgrading in emerging market societies. *Academy of Management Journal* 52(6): 1270–1296.

McDonald F., Q. Huang, D. Tsagdis, *et al*. 2007. Is There Evidence to Support Porter-type Cluster Policies? *Regional Studies* 41.1: 39–49.

McEvily, B., A. Zaheer. 1999. Bridging ties: a source of firm heterogeneity in competitive capabilities. *Strategic Management Journal* 20(12): 1133–1156.

Mesquita, LF, SG Lazzarini. 2008. Horizontal and vertical relationships in developing economies: implications for SMEs' access to global markets. *Academy of Management Journal* 51(2): 359–380.

Multilateral Investment Fund (MIF). 2007. *MIF cluster action plan. Promoting local competitiveness*. MIF/GN-124, 26 June. Washington, D.C.: Inter-American Development Bank.

Moreno, JL. 1934. *Who shall survive?* Beacon, NY: Beacon House.

Mori, J., Y. Kajikawa, I. Sakata. 2010. *Evaluating the Impacts of Regional Cluster Policies using Network Analysis*. International Association for Management of Technology IAMOT 2010 Proceedings.

Newman, MEJ. 2001. The structure of scientific collaboration networks. *Proceedings National Academy Science USA* 98: 404–9.

----. 2003. The structure and function of complex networks. *SIAM Review* 45(2): 167–256.

Nishimura J., Okamuro H., 2011. Subsidy and networking: The effects of direct and indirect support programs of the cluster policy, *Research Policy* 40: 714–727.

Oldsman, E., K. Hallberg. 2002. *Framework for Evaluating the Impact of Small Enterprise Initiatives*. SED Working Paper No. 3, Swiss Agency for Development and Cooperation.

Pacheco, M. 2010. *Politicas publicas para el fortalecimiento de clusters: La experiencia BID en Brazil y America Latina*. V Congreso Latinoamericano de Clusters, Medellin, 24 June 2010.

Paruchuri, S. 2010. Intraorganizational Networks, Interorganizational Networks, and the Impact of Central Inventors: A Longitudinal Study of Pharmaceutical Firms. *Organization Science* 21(1): 63–80.

Peres, W. 2004. *Políticas sectoriales y de desarrollo de clusters en América Latina y el Caribe, Políticas sectoriales y de desarrollo de clusters en América Latina y el Caribe*. Santiago de Chile: CEPAL.

Pietrobelli, C., R. Rabellotti. 2007. *Upgrading to Compete. Clusters and Value Chains in Latin America*. Cambridge, MA.: Harvard University Press.

Pietrobelli C., R. Rabellotti. 2007. Business Development Service centres in Italy: close to firms, far from innovation. *World Review of Science, Technology and Sustainable Development* 4(1): 38–55.

Pietrobelli C., L. van Oyen, J. Dobinger. 2009. *Independent Thematic Evaluation of the UNIDO Cluster and Networking Development Initiatives*. Vienna: United Nations Industrial Development Organization.

Podolny, JM 1993. A status-based model of market competition. *American Journal of Sociology* 98(4): 829–872.

Porter, ME. 1998. *On Competition*. Boston: Harvard Business School Press.

Powell, WW. 1990. Neither market nor hierarchy: network forms of organization. *Research in Organizational Behavior* 12: 295–336.

Powell, WW, KW Koput, L. Smith-Doerr. 1996. Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology. *Administrative Science Quarterly* 41: 106–45.

Quiroz J.A. 2007. Cluster Economics: Elements for Program Evaluation, IADB internal document.

Raines, P. 2002. *The Challenge of Evaluating Cluster Behaviour in Economic Development Policy*. Paper presented to the International RSA Conference: Evaluation and EU regional policy: New questions and challenges.

----. 2001. *The cluster approach and the dynamics of regional policy-making*. Regional and Industrial Policy Research Paper 47, European Policy Research Centre, University of Strathclyde, U.K.

Robins, GL., PE Pattison, S. Wasserman. 1999. Logit Models and Logistic Regressions for Social Networks: III. Valued Relations. *Psychometrika* 64: 371–94.

Robins, GL., PE Pattison, Y. Kalisha, *et al*. 2007. An introduction to exponential random graph (*p**) models for social networks. *Social Networks* 29(2): 173–191.

Rosenfeld, SA. 2002. *Creating Smart Systems A guide to cluster strategies in less favoured regions*. European Union-Regional Innovation Strategies.

Rowley, T., D. Behrens, D Krackhardt. 2000. Redundant governance structures: An analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic Management Journal* 21(3): 369–386.

Russo, M., F. Rossi. 2009. Cooperation Networks and Innovation: A Complex Systems Perspective to the Analysis and Evaluation of a Regional Innovation Policy Programme. *Evaluation* 15(1): 75–99.

Schmiedeberg, C. 2010. Evaluation of Cluster Policy: A Methodological Overview. *Evaluation* 16(4): 389–412.

Schmitz, H. 1995. Collective Efficiency: growth path for small-scale industry. *Journal of Development Studies* 31: 529–66.

Serrat, O. 2009. *Social Network Analysis*. Knowledge Solutions N° 28, Asian Development Bank, Manila.

Snijders, TAB. 2005. Models for longitudinal network data. In *Models and Methods in Social Network Analysis*, eds. P. Carrington, J. Scott, S. Wasserman. New York: Cambridge University Press.

----. 2001. The statistical evaluation of social network dynamics. *Sociological Methodolog*y: 361–395.

Snijders, TAB., EP Pattison, GL Robins, *et al*. 2006. New Specifications for Exponential Random Graph Models. *Sociological Methodology* 36: 99–153.

Snijders T.A.B., Van De Bunt G.G., Steglich C. E. G. 2010. Introduction to Stochastic Actor-Based Models for Network Dynamics. *Social Networks* 32: 44–60.

Staber U. 2001. The Structure of Networks in Industrial Districts, *International Journal of Urban and Regional Research*, 25: 537-552.

Ubfal, D., A. Maffioli. 2010. *The impact of funding on research collaboration : evidence from Argentina*. IDB Working Papers No. IDB-WP-224, Washington, D.C.

Uzzi, B. 1997. Social structure and competition in interfirm networks: the paradox of embeddedness. *Administrative Science Quarterly* 42: 35–67.

Uzzi, B., J. Spiro, D. Delis. 2002. *Emergence: The Origin and Evolution of Ccareer Networks in the Broadway Musical Industry, 1877 to 1995*. Working Paper, Northwestern University, Evanston, IL.

Wasserman, S., K. Faust. 1994. *Social Network Analysis. Methods and Applications*. Cambridge, MA.: Cambridge University Press.

Watts, DJ. 2004. The "New" Science of Networks. *Annual Review of Sociology* 30: 243–270.

Watts, DJ, SH Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature* 393 (6684): 409–10.

Winters, P., L. Salazar, A. Maffioli. 2010. *Designing Impact Evaluations for Agricultural Projects*. IDB Impact-Evaluation Guidelines, Technical Notes No. IDB-TN-198, Washington, D.C.

Zaheer, A., GG Bell. 2005. Benefiting from network position: Firm capabilities, structural holes, and performance. *Strategic Management Journal* 26: 809–825.

APPENDIX

A Basic Glossary for Social Network Analysis

1. A network is a finite set or sets of **actors** and the relations defined on them.
2. An **actor** is a discrete individual, corporation, or other organizational unit (e.g., people, firms, university departments, government agencies). In graph theoretical terminology, each actor is identified as a **node** (or **vertex**) in the network.

3. Actors are linked to each other by social ties. Each network consists of connections between nodes based on the **same type** of social tie. Social ties may be of any type, ranging from friendship between two individuals to economic transactions, an exchange or transfer of knowledge, labor mobility, etc. In graph theoretical terminology, each tie is identified as an **edge** (or **line,** or **arc** if directional). Edges can be **directional** when they express a flow from one node to another (e.g., transfer of money), and **nondirectional** when relationships are bi-directional by nature (e.g., the geographical distance between two actors). They can also be **dichotomous** (on a scale of 0 to 1) and **valued** (weighted on the basis of value, importance, frequency, etc.).
4. An important distinction is between a **one-mode** network and a **two-mode** network. The former is composed of actors that are of the same type (e.g., all nodes are business firms) and accounts for the linkages existing between these actors. A two-mode network is composed by **two distinct** sets of actors (e.g., business firms and universities), and a two-mode network dataset contains measurements on which actors of one type (say, firms) have ties to actors of another type (universities).
5. Network data represented in two-way matrices are called sociomatrices or **adjacency matrices**.
6. A **path** is a sequence of nodes and lines in which all nodes and lines are distinct—no actor is part of the path more than once.
7. The shortest path between two nodes is referred to as **geodesic** (or **distance**).
8. The **diameter** of a network is the length of the largest geodesic between any pair of nodes.
9. The **component** of a graph is a maximal connected subgraph. This means that it is a subgraph in which there is a path between all pairs of nodes and the subgraph (all pairs of nodes are reachable), and there is no path between a node in the component and any node not in the component (which means that it is **maximal**).
10. A maximal **complete** subgraph is a component where all lines are present and all nodes are adjacent. Being adjacent means being connected by an edge.

Key websites and information sources on SNA

International Network for Social Network Analysis (INSNA): http://www.insna.org/

Software for Social Network Analysis: http://www.analytictech.com/

http://vlado.fmf.uni-lj.si/pub/networks/pajek/

Software and courses for Stochastic Network Analysis: http://www.stats.ox.ac.uk/~snijders/siena/

SNA publication in Spanish: http://revista-redes.rediris.es/

Latin American Meetings on Social Network Analysis: http://reunionredes2011.wordpress.com/

http://reunionredes2011.wordpress.com/antecedentes/