Paper to be presented at DRUID17  
NYU Stern School of Business, New York, June 12-14, 2017

**Knowledge base decomposability: the role of alliances in the industry-wide network**

Mario Vitale  
LUISS Guido Carli University  
Department of Business and Management  
mvitale@luiss.it

Federica Brunetta  
LUISS Guido Carli University  
Department of Business and Management  
fbrunetta@luiss.it

**Abstract**

The process of generating innovations is not only a matter of recombining knowledge elements: it implies navigating in the vast space of all possible combinations. In navigating this space, firms are limited by their ability to process all the potentially relevant variables so they rely on their knowledge bases. These knowledge bases are the result of firms’ choices to create couplings among knowledge elements. These coupling choices act as a guide to decompose the search space. This paper focuses on the role played by external environment, and more specifically, by alliances in the industry wide network in influencing these coupling choices and accordingly, the emergence of different knowledge bases. Different knowledge bases are associated with different knowledge structures, each possessing diverse levels of innovation propensity and rigidity towards change. This paper argues that specific positions in the industry wide network are related to different knowledge bases possessing different degrees of rigidity to change. More specifically, centrality and structural holes spanning can favor or prevent the recombination in the knowledge base toward an optimal degree, reducing or increasing this rigidity. Indeed, through the information exchange and knowledge sharing, firms can discover better configuration of knowledge bases and this may activate recombination. Research hypotheses are developed and investigated within worldwide nanotechnology industry. Data are driven from Patstat, S&P Capital IQ and Thomson Reuters Datastream for 78 firms over a time window of 10 years.

Jelcodes:O32.O30
Knowledge base decomposability:

the role of alliances in the industry-wide network

Abstract

The process of generating innovations is not only a matter of recombining knowledge elements: it implies navigating in the vast space of all possible combinations. In navigating this space, firms are limited by their ability to process all the potentially relevant variables so they rely on their knowledge bases. These knowledge bases are the result of firms’ choices to create couplings among knowledge elements. These coupling choices act as a guide to decompose the search space. This paper focuses on the role played by external environment, and more specifically, by alliances in the industry wide network in influencing these coupling choices and accordingly, the emergence of different knowledge bases. Different knowledge bases are associated with different knowledge structures, each possessing diverse levels of innovation propensity and rigidity towards change. This paper argues that specific positions in the industry wide network are related to different knowledge bases possessing different degrees of rigidity to change. More specifically, centrality and structural holes spanning can favor or prevent the recombination in the knowledge base toward an optimal degree, reducing or increasing this rigidity. Indeed, through the information exchange and knowledge sharing, firms can discover better configuration of knowledge bases and this may activate recombination. Research hypotheses are developed and investigated within worldwide nanotechnology industry. Data are driven from Patstat, S&P Capital IQ and Thomson Reuters Datastream for 78 firms over a time window of 10 years.
Introduction

Innovation can be generated either from a novel combination of technological components (Nelson and Winter, 1982; Basalla, 1988), or through their reconfiguration (Henderson and Clark, 1990). However, the process of generating innovations should not merely be seen as matter of recombining knowledge elements: it relates to finding a way to navigate all possible combinations within a vast space (Fleming and Sorenson, 2001). In navigating this space, firms are limited by their ability to process all the relevant variables, as well as their intricate set of interactions. Firms may not be able to conceive the full set of combinations or to determine ex ante the causal linkage between possible combinations and possible outcomes (Gavetti and Levinthal, 2000). Therefore, firms rely on their cognitive representations of the world to decompose the search space. These representations drive firms in the interpretation and recombination of the knowledge base, by detecting which knowledge components have the highest potential in being matched and shall, therefore, be combined, and on the other hand which ones should be kept apart, solving in this way the so-called “combinatorial explosion problem”, that represents the difficulty in analyzing in depth all possible combinations.

Since organizations do not know ex ante the real search space and all the interdependencies among knowledge elements that compose it, they are forced to guess on them. These guesses generate a specific structure of the knowledge base of an organization and different structures have important effects on the ability of generating useful innovations. However, as firms learn and acknowledge the real interdependencies, they can change their couplings. Change in the knowledge structures is important form of adaptation. Different coupling patterns determine the knowledge base of a firm and assign it a structure with different degree of “decomposability”, from completely modular to completely integrated, making these changes more or less feasible (Yavavaram and Ahuja, 2008).

We focus on the role played by external environment, and more specifically, by alliances in the industry wide network in influencing decisions on couplings and ultimately their changes. Different knowledge structures are associated with different levels of rigidity towards change (Yavavaram and Ahuja, 2008). We argue that specific positions in the industry-wide network can favor or prevent the recombination in the knowledge base toward an optimal degree, reducing or increasing this rigidity. Indeed, through the information exchange and knowledge sharing, firms can discover better configuration of knowledge bases and this may activate recombination.
Alliances in the industrywide network shape the interpretation of knowledge, especially when evaluating which elements and components are relevant and deserve attention. Previous studies, taking into account the network of connections among firms and their innovations in technological landscapes, have highlighted the subjectivity of technological developments and rejected the idea that technologies are deterministic and that superior innovations always dominate inferior ones (Hughes, 1987). These studies suggest that firms decide which technological options to pursue looking at the decision of other firms. In this sense, technological change becomes the result of interdependences among communities of firms (Rosenkopf and Tushman, 1998; Tushman and Rosenkopf, 1992). Networking allows firms to effectively explore new knowledge. Actors bring experience and knowledge accrued through interaction with other partners into their relationships with the focal organization (Gulati and Gargiulo, 1999). Nonetheless, internal capabilities and external collaboration jointly facilitate the exploitation of existing knowledge (Mowery 1989).

In high-growth, technology-intensive industries such as nanotechnology, networks are created among actors characterized by different knowledge, competencies, access to assets and information (Pisano, 1991; Powell et al., 1996; Barley et al., 1992; Walker et al., 1997; Arora and Gambardella, 1994). Actors may include competitors, collaborators and stakeholders. We focus on networks generated by strategic alliances, as these are recognized as one of the way that firms use to craft their frames and to navigate the search space, reducing complexity and uncertainty. Alliances imply exchanging, sharing, or co-developing products, services or technologies (Gulati, 1998). Alliance networks can be considered as organizational devices generating heterogeneous learning processes (Riccaboni and Moliterni, 2009), useful to guide organizations in the search space.

We develop research hypotheses and investigate them within worldwide nanotechnology industry, which appears particularly suitable for our analysis due both to its high R&D intensity and sustained pace of innovation, and the fact that its actors frequently engage in strategic alliances. We gathered knowledge base data looking at technology classes, patents and patent citations. Data on alliances are built to catch all the relevant alliances in the industry-wide network.
Theoretical background

As boundedly rational actors we are unable to elaborate perfect and exhaustive cognitive representations to describe the world (Thagard, 1996). Representations are used in order to describe the world, but also to provide a simplification of the complexity of spatial (Porac et al., 1989), as well as causal or temporal relationships (Weick, 1979) and to ease the understanding of the interaction among choices and actors. Despite these limitations, cognitive representations represent one of the most important drivers of managerial actions and choices (Tversky and Kahneman, 1986) and very often a firm’s strategy is based on managers’ representations of their problem spaces (Simon, 1991).

Cognitive representations result from historical experiences rather than being dependent on how the knowledge of the environment is currently developed (Tripsas and Gavetti, 2000). For these reasons, in high velocity environments, individuals fail to adapt their mental models, changing their beliefs and this results in scarce organizational performance (Brown and Eisenhardt, 1998; Barr et al., 1992).

However, it is worth noting that cognitive representations work as a templates that guide the choice, but do not constrain, the set of behaviors that actually arises (Gavetti and Levinthal, 2000). For instance, within a firm the existing set of routines may contribute to act as starting point for choices, although the template does not specify the latters. Otherwise, reliance on past experience might be the rule (Gavetti and Levinthal, 2000).

Cognitive representations play a key role in driving the process of re-combinative search, offering a guide to decompose the search space and to create specific mapping of couplings. In other words, relying upon its cognitive representation of the world, a firm performs the decomposition of the search space for innovations, deciding which elements of knowledge have the potential to jointly work well and have to be combined and which ones should not. Thus, cognitive representations have a clear impact on a firm’s knowledge base structure. Coupling choices reflect, at the same time, the implicit and the explicit organizational assumptions – the so-called “best guesses” (Yayavaram and Ahuja, 2008; p. 377) - regarding the interdependencies between its cognitive map of the world and the knowledge components. In this sense a given mapping can be more or less faithfully representative of the actual environment in which the organization operates. This implies that decision on coupling may differ across distinct entities despite the fact that everyone has a common set of underlying interdependencies.
A firm’s cognitive representation tends to reside in its communication patterns (Henderson and Clark, 1990), routines (Nelson and Winter, 1982), organizational structure and beliefs. These information patterns, in any firm, impact on the understanding of knowledge, and influence decisions on the important elements and which couplings seem to be the most likely. Within the firm, informal (Moorman and Miner, 1997) or formal routines (e.g. standard operating procedures) suggest the knowledge elements that should be used in the search process. For instance, firms can have formal routines to transfer knowledge across units and for acquiring knowledge from outside. Similarly, it is possible to find formal patterns of communication among researchers or across R&D units (Allen, 1997), determining which ones may eventually interact with each other.

The same organizational structure with its formal and informal communication networks is demonstrated to play an important role in determining cognitive representations. As routines, organization structure, and beliefs vary across firms, we observe differences in their choices of couplings.

However, we focus on another important element that prior literature identifies as key in the process of search for innovation: the role played by the network of strategic alliances. As a matter of fact, cognitive representations emerge from the network of connections among firms and their innovations in technological landscapes. Prior studies suggest that innovations are the result of interdependencies among communities of firms (Rosenkopf and Tushman, 1998; Tushman and Rosenkopf, 1992). Indeed, it is demonstrated that the innovation efforts performed by other firms influences a firm’s innovation efforts (Rosenkopf and Nerkar, 1999). Moreover, a firm is influenced by the beliefs shared across the entire industry about perceived interdependencies (Yavavaram and Ahuja, 2008). Thus, the relational context in which firms are embedded seems to play a key role in the emergence of specific coupling patterns.

**The role of alliances**

Since firms do not know ex-ante the real search space and all the interdependencies among knowledge elements that compose it, they are forced to guess on them. These “guesses” are the result of their cognitive frames and we define them as couplings, namely the “result of search-related decisions of an organization on how knowledge elements should be recombined” (Yavavaram and Ahuja, 2008: 344). Couplings generate a firm’s knowledge base,
determining different knowledge structures, each with different degree of “decomposability”, from completely modular to completely integrated (Yavavaram and Ahuja, 2008), with an “optimal” degree of decomposability seeming to lie between the extreme cases. These types of couplings vary across distinct entities although the set of underlying interdependencies among knowledge elements might be common to everyone.

The coupling pattern of each firm can be updated as firms learn and acknowledge new interdependencies. The discovery of new interdependencies can be the result of several factors. Among them, previous studies have analyzed scientific knowledge (Rosenberg, 1982; Klevorick et al, 1995), suppliers, users (von Hippel, 1988), alliances (Ahuja, 2000; Ahuja and Katila, 2001), and mobility of inventors. Here we focus on alliances. Indeed, alliances can influence the emergence of specific structure by learning either vicariously (March, et al. 1991) or at the population-level (Miner and Haunschild, 1995), from the representations of successful actors. Firms may set their couplings in the attempt to imitate their competitors or to acquire knowledge from outside. Moreover, as a consequence of shifts in technology of an industry (i.e. radical innovation/disruptive innovation), firms may be forced to change their coupling pattern in the attempt to capture the characteristics of the new search space.

Following scholars who have well underlined that knowledge construction happens at the social level (Mizruchi and Fein, 1999), we argue that different portfolios of strategic alliances impact on the internal knowledge base creation following two different mechanisms. First, different templates can be generated by external influences (Scott, 1995; Goffman, 1983); each of them, crafts internal cognition and thus the choice of elements to combine. Second, through partnerships, organizations pursue externally driven search strategies. In this light, we recognize the role of networks in disseminating ideas and information, allowing knowledge access and enabling novel combination (Powell et al. 1996, 2005).

Networking allows organizations to effectively explore new knowledge. Actors bring experience and knowledge accrued through interaction with other partners into their relationships with the focal organization (Gulati and Gargiulo, 1999). Nonetheless, internal capabilities and external collaboration jointly facilitate the exploitation of existing knowledge (Mowery 1989).

In high-growth, technology-intensive industries such as nanotechnology, networks are created among actors characterized by different knowledge, competencies, access to assets and information (Pisano, 1991; Powell et al., 1996; Barley et al., 1992; Walker et al., 1997; Arora and Gambardella, 1994). Actors may include competitors, collaborators and stakeholders. We
focus on networks generated by strategic alliances, as these are recognized as one of the way that firms use to craft their frames and to navigate the search space, reducing complexity and uncertainty. Alliances imply exchanging, sharing, or co-developing products, services or technologies (Gulati, 1998). Alliance networks can be considered as organizational devices generating heterogeneous learning processes (Riccaboni and Moliterni, 2009) useful to guide organizations in the search space.

We cannot neglect the role played by strategic alliances in both influencing an organization’s cognitive frame and accruing and shaping its knowledge base. External collaboration offers complementarities to internal knowledge, which affect coupling decisions. It is also relevant to consider that the position an actor takes within a network is crucial, as it influences its potential outcomes, the possible knowledge access and its behavior (Powell et al., 1996; Walker, et al., 1997). Prominent actors, defined by network literature as those with an higher degree of centrality, have more immediate access to new and important knowledge, trends in cognitive frame and intellectual fashions (Perrow, 1986; Mizrouchi and Fein, 1999).

**Hypotheses development**

Prior studies suggest that networks fuel an actor’s learning capability and internal expertise (Powell et al, 1996). Networks also generate positive effects on the performance, as they provide channels though which partners can share experiences and knowledge (DeBresson and Amesse, 1991; Freeman, 1991; Gulati and Gargiulo). More specifically, the position that a firm can assume within the network influences its potential for opportunity scanning, competence enhancement and reach (Powell et al., 1996). Among the different positions that a firm can occupy in a network there is centrality. Since its initial formulation (Bavelas, 1948), centrality has been largely studied and many scholars have analyzed advantages and disadvantages associated to it. Central positions motivate an actor’s capability to influence or be privileged (Freeman, 1979), to be in a prominent location or be more visible to others (Wasserman and Galaskievicz, 1994), to retain or obtain power (Brass and Burkhardt, 1993) or, finally, achieve higher status or prestige (Bonacich, 1972).

Summing up, on of the core benefits of centrality is that of obtaining superior information and, thus, generate an advantage (Rowley, 1997). So, possessing an highly central position is related to receiving more information and resources, and fastly. Sharing of information within the network can happen following different mechanisms, that can be classified on the basis of
type of trajectory that the information flow may take (geodesics, paths, trails, or walks) and the mechanisms through which information spreads (broadcast, serial replication, or transfer) (see Borgatti, 2005 for an in-depth explanation). When information moves via unrestricted walks and each node impacts on all of its neighbors at the same time, the central actors have access to the best information (and, at the same time, influence information generation).

It is important to underline that centrality can be expressed by different measures, each accounting for different elements flowing through network ties. Thus, the relevance of a node’s position cannot be defined without understanding how information flows through ties.

Here we follow Borgatti (2005) referring to Eigenvector centrality. Eigenvector centrality measures how much a node is adjacent to nodes that are themselves highly central.

The firm’s centrality influences its capacity of acting as an information hub that disseminates and receives knowledge and information across the network. A firm in this position can be strategically important, and play a key role: central actors may enhance the mobility of knowledge, equally distribute value, and foster trust within partners; finally, they can be crucial in promoting stability within the network (Dhanaraj and Parkhe, 2006).

Since knowledge is broadly distributed among different firms in the search space, the network of firms is the locus of innovation (Powell and Brantley, 1992) and can play a key role in affecting the emergence of knowledge bases. In a rapidly growing field characterized by an intense activity of R&D, to stay current a firm needs a hand in the search process. For this reason, in settings in which knowing-how is pivotal, firms should be able to combine both in-house (individual learning) and cooperative search (social learning).

Centrality enables firms to learn more extensively from others. However complementarity, not substitution, exist between internal capability and external collaboration (Mowery and Rosenberg, 1989; Arora and Gambardella, 1994). Networks give access to information and knowledge that is difficult to generate internally (Nelson, 1990). More specifically, networks serve as loci of innovation allowing firms to access knowledge on the real interdependencies that exist in the search space that would be unavailable otherwise, and at the same time, test internal assumptions on couplings. Being central in the network enables firms to cope with the rhythm and progress of high-potential technological developments. When the sources of knowledge are disparate, collaboration helps firms to understand that they need to access ideas and information from different sources. Finally, when an actor is connected to central actors, the information is higher.
For this reason, we argue that central firms have a higher probability to access high quality information on the real interdependencies in the search space. Central firms can, in fact, benefit from expertise and trial-and-error process undertaken by other firms that are in the network. In this sense firms in a central position may learn faster the right coupling pattern to reach a higher innovative performance. Finally, being connected and central crafts the reputation of a firm and influences its visibility, allowing access to resources. In sum, central firms are able to assess and evaluate different coupling patterns, selecting the better combination of knowledge elements and moving toward it. In other words, we expect central firms to decrease their distance from the optimal degree of decomposability.

Thus, we hypothesize that:

HP1: The higher the firm’s centrality in the industry-wide network of alliances, the less the distance of its knowledge bases form the optimal decomposability level.

Actors occupying central positions with respect to closeness usually tend to be more productive in communicating information to the other actors (Wasserman & Faust, 1994; Beauchamp, 1965), and are not dependent upon others as intermediaries (Freeman, 1979). As pointed out by Borgatti (2005) these actors are positioned to obtain novel information early. A central actor is able to move information across the network by using a lower number of intermediaries.

Nevertheless, when looking at these specific aspects, identified by closeness centrality, factors such as strategic orientation and external constraints may influence the potential outcomes of network position. Broadly speaking, a firm’s closeness centrality overlooks how well the actor’s partners are connected. Firms with a high level of closeness can quickly transmit information and influence others. Closeness is based on the ideas of efficiency and independence (Freidkin, 1991). Being close to others in the network means being more efficient in transmitting information. Moreover, firms high on closeness have independence in the sense that they do not need to seek information from other more peripheral actors. However, two negative consequences may arise: first, an actor which is “close” to a large number of alters can be overloaded or eventually strongly depend upon its connections; second, ties represent sources of resources and information, but at the same time a gate, or “weak” point through which resources and knowledge may drain (Gnyawali & Madhavan, 2001; Brunetta et al., 2015). These two elements could hinder the process of information selection driving firms in
establishing a coupling pattern that is not optimal. Moreover, as closeness centrality increases, actors may face an increase in the costs of coordination (Gulati and Singh, 1998; Gargiulo and Benassi, 2000), as well as cognitive problems (Hansen et al., 2005), with a subsequent reduction of the benefits of centrality. Being visible may constrain the ability of an actor to process new ideas or information, and also to be as flexible as others, which are less constrained. Finally, firms with high closeness toil to access to knowledge produced in the periphery of networks. This could harm the ability of firms to perform distant search, that we have seen is fundamental to change knowledge structures. Thus, we hypothesize:

**HP2: The higher the firm’s closeness centrality in the industry-wide network of alliances, the more the distance of its knowledge bases form the optimal decomposability level**

Another important element to take into account when we discuss on the role of network positions as driver for innovative performance is structural holes. Prior literature emphasized the idea that structures that are rich in structural holes might provide an actor with privileged, efficient, or timely access to resources and broader, more diverse information (Emerson, 1962; Burt, 2001; Ahuja, 2000; Rowley et al., 2000). They are therefore viewed as capable of creating fertile ground for the achievement of innovation-oriented tasks and for encouraging entrepreneurial activities (Bower, 1970; Burgelman, 1983; Walker et al., 1997). The “dark side” of structural holes is related to the “action problem” (Obstfeld, 2005), which reflects an inability to effectively coordinate remote participants.

Access to a broader range of information promotes the timely gathering of information and increases actors’ chances of identifying opportunities (Burt, 1997, 2004; Powell et al., 1996). Notwithstanding these benefits, networks rich in structural holes can experience an obstacle-related execution problem when attempting to coordinate or implement fresh ideas (Obstfeld, 2005). Actors that are weakly related tend to present strong dissimilarities that hinder the potential to integrate the diverse array of available knowledge (Tiwana, 2008). So, if one hand structural holes could be beneficial to access to fresh knowledge, on the other hand they offer less information about the experienced coupling of patterns in this industry-wide network. Without accessing this knowledge, it becomes harder for firm to evaluate and select the optimal structure of the knowledge base. Knowing the coupling patterns performed by other
firms is important for the firm in order to update its “guesses” on the real interdependencies in the search space (Yayavaram and Ahuja, 2008; Author citation, 2017).

Thus, we hypothesize that:

**HP3:** The higher the firm’s number of structural holes in the industry-wide network of alliances, the more the distance of its knowledge bases form the optimal decomposability level

**Methods**

**Empirical setting and data**

The empirical setting of our study is the worldwide nanotechnology industry. Nanotechnology is considered to be one of the key technologies of the 21st century with has reached a market volume of EUR 1 trillion in 2015. This setting is characterized by a high R&D intensity since technology search plays a key role in this industry, and by an intense activity of strategic alliances, due to an well functioning market for ideas (Rotheamel and Thursby, 2007).

Data to calculate knowledge bases and their degree of decomposability are based on technology classes, patents and patent citations. These data were drawn by PATSTAT. PATSTAT is a Worldwide Patent Statistical Database developed by European Patent Office. The PATSTAT dataset includes 20 tables with bibliographic data, citations and family links of about 70 million applications (in 80 countries). More specifically, we extracted all patents in the Nanotechnology industry and their relative information following the International Patent Classification system (IPC) and the Y01N label and B82 labels, due to the fact that beginning 1 January 2011 the EPO has introduced a new class symbol (B82) to classify nanotechnology, replacing the label Y01N.

Network data were elaborated starting from data on alliances among firms in our patent database. In this sense we have built a database on the alliances in the nanotechnology industry. Data on alliances were gathered by the S&P Capital IQ database. Originally designed to address the needs of the investment banking community, Capital IQ focuses on financial information and analysis for companies and industries. Overall, S&P Capital IQ database includes more than 5000 data items and industry-related metrics on different industries.

In order to identify alliances and network relations, we used the database function “Key Developments” to screen announcements on alliances and transactions based on description, date, type, status. Specifically, we focused on strategic alliances performed from 2001 to 2013.
Once the partners in each alliance were identified, network data were collected in matrices “actor per actor”, which identify the ties between the actors weather a strategic alliance is present or not. More specifically we have inserted “1” in the case in which two actors are connected by a strategic alliance and “0” when this is not the case. Through the UCINET software 6.0 (Borgatti et al., 2002), we created primary affiliations data to calculate specific ego network indicators, to evaluate point centrality - eigenvector and closeness - as well as a measure of structural holes - for each actor analyzed. Network data were computed on the whole network of actors involved in the alliances, not only among selected companies, in order to have a full description of network relations.

We obtained control variables from Thomson Reuters Datastream Database, a global financial platform covering equities, stock market indices, currencies, company fundamentals, and key economic indicators. Datastream is an economic and financial time series database. We collected information on firm performance and firm size. Data recorded in different currencies were converted to US dollars using the yearly average exchange rates (for each of the years analyzed). Data were cross-checked with financial information contained in the S&P Capital IQ. We identified 78 nanotech firms across all over the world. Information on subsidiaries came from S&P Capital IQ database.

**Variables**

**Independent Variables**

Centrality measures have first been studied by Bavelas (1948), who tested the hypothesis that central positions could confer influence to the actors; the concept of centrality has since then been applied in a variety of analysis, confirming that centrally positioned actors are more influential and occupy a privileged position. There is an extensive literature on centrality, which highlights a variety of measures and the importance of centrality as a structural attribute of social networks. Nevertheless, each measure is best used to describe and evaluate specific network attributes, especially as different types of “traffic” flow through network ties (Borgatti, 2005). Thus, we focused on two main measures of centrality, eigenvector and closeness.

**Eigenvector Centrality** This measure has been proposed by Phillip Bonacich (1972; 1987) as a modification of the degree centrality approach. It is based on the idea that the centrality of an
ego is based on the summed connections to the alters, weighted by their centralities (Bonacich, 1972). Basically, this measure expresses the power of the actor: central egos tight to other central egos are more “powerful” than central egos tight to peripheral egos. Determining this type of centrality creates a vicious circle (Degene and Forsé, 1999) as one would need to determine which centrality value an actor i assumes knowing the centrality of all nodes j that connect to i, that ultimately depend on the centrality of i itself. The formal notation for this is:

\[ C_i = \sum_j r_{ij} C_j \]

in which \( r_{ij} \) is the value of the i-j relation. This problem can be resolved by an algorithm (Bonacich, 1972): being \( R \) the matrix of relations, \( \lambda \) a constant used to avoid zero solutions (therefore, the eigenvalue of \( R \)) and \( e \) the vector of centrality scores for each node (the eigenvector or \( R \)) it can be expressed as:

\[ \lambda C_i = \sum_j r_{ij} C_j \]

that in matrix notation becomes:

\[ \lambda e = R e \]

We can obtain the centrality indices for all the actors in the network by choosing the weighted eigenvector that matches the greatest eigenvalue in the matrix.

**Closeness centrality.** This measure is based upon “the minimum cost or time for communicating with all other points.” (Freeman, 1979: 225). In other words, it is based on distance and takes into consideration not only the connections to immediate alters, but the closeness to all network actors.

The underlying assumption in the calculation of this index is that central actors can interact with others and do not rely on others for their information, as they are placed on the communication paths linking pairs of other actors (Bavelas, 1948; Shaw, 1954). Moreover, these actors are in the position of withholding or distorting information in transmission (Freeman, 1979). Sabidussi (1966) has quantified the measure of closeness as a function of its geodesic distances to all points in the graph; this means that as the geodesic distances increase in length, the centrality decreases. Following Freeman (1979) and given \( d(pi, pk) \) = the number of edges in the geodesic linking pi and pk, we get Sabidussi’s measure of decentrality of \( p_i \) as
\[ C_c(p_k)^{-1} = \sum_{i=1}^{n} d(p_i, p_k). \]

This measure increases as the closeness between \( p_k \) and other points decreases (distance grows), therefore it is inversely related to the centrality of point \( p_k \). Nonetheless, as this measure is dependent upon total nodes within the network, Beauchamp (1965) formulated another measure of relative point centrality as:

\[ C_c(p_k) = \left[ \frac{\sum_{i=1}^{n} d(p_i, p_k)}{n-1} \right]^{-1} = \frac{1}{n-1} \sum_{i=1}^{n} d(p_i, p_k). \]

This represents the inverse measure of the average distance between \( p_k \) and other nodes, thus, the inverse of the ratio by which \( p_k \) exceeds its minimum distance. Thus, \( C_c(p_k) \) measures directly point centrality as distance-based. The measures \( C_c(p_k)^{-1} \) and \( C_c(p_k) \) are both closeness-indexes of centrality.

**Structural holes.** This variable measures the potential to access diverse knowledge and information through the brokerage role played by central nodes in network regions rich with holes. It was computed starting from the constraint index. The constraint measures the extent to which ego is invested in people who are invested in others among ego’s alters. The opportunities for actor \( i \) are constrained by \( j \) if another contact, \( q \), that is linked to \( i \)’s network has invested heavily in a relationship with actor \( j \). This relation is expressed by

\[ p_{ij}p_{qj}, \]

where \( p_{qj} \) represents the proportional strength of \( q \)’s relation with \( j \). The higher this relation, the higher the investment in contacts between \( i \) and \( j \), and the lower the opportunities to develop a structural hole among them. If we then consider all contacts \( q \), and add the direct connection between \( i \) and \( j \) (\( p_{ij} \)), we can define constraint as the proportion of \( i \)’s network that directly or indirectly involves \( j \) as:

\[ p_{ij} + \sum_{q \neq i,j} p_{iq}p_{qj} \]

in which there are still indirect relations between \( i \) and \( j \) even if the direct relation \( p_{ij} \) is subtracted. The expression also measures the lack of structural holes with which actor \( i \) could negotiate any upcoming demand from \( j \). It is then necessary to multiply the product of investment with the lack of structural holes (i.e., to square the expression) in order to obtain a
measure of the constraint on actor $i$ stemming from the lack of primary holes around $j$. In this case, the constraint $c_{ij}$ can be calculated as:

$$c_{ij} = \left( p_{ij} + \sum_{q} p_{iq}p_{qj} \right)^2, q \neq i, j$$

Constraints range from a minimum of $p_{ij}$ (squared), in which the actor $j$ is disconnected from all other contacts, to a maximum of 1, in which $j$ is $i$’s only contact. At lower values of this index, a firm’s ego-network is rich in structural holes. The sum of the above constraints across contacts $j$ provides a measure of the aggregate constraint on $i$’s opportunities within the network. Accordingly, the structural holes’ measure is the complement to 1 of the constraint:

$$StructuralHoles = \sum 1 - c_{ij} = \sum 1 - \left( p_{ij} + \sum_{q} p_{iq}p_{qj} \right)^2$$

**Dependent variables**

The variable distance from optimal degree of decomposability measures how far the degree of decomposability of a firm’s knowledge base is from the optimal degree of decomposability that allows the maximal innovative performance. This variable is calculated by taking the norm (i.e., the absolute value of the difference) of the optimal degree of decomposability and the actual degree of decomposability of a firm’s knowledge base, with the latter calculated as the clustering coefficient (Watts and Strogatz, 1998) of the network whose nodes are the technology classes. We follow Yayavaram and Ahuja (2008), defining the clustering coefficient for a node (technological class) with $k_i$ ties as $CCi = n_i / [k_i * (k_i – 1) / 2]$, where $n_i$ is the total amount of ties (citations) between the $k_i$ neighbors of node $i$. The denominator is the maximum possible number of ties between the $k_i$ neighbors of node $i$, and the numerator is the actual number of ties that exist. The clustering coefficient for the network is $CCi$ averaged over all nodes.

**Control Variables**

We include several control variables in our models. More specifically, at the patent level, we control for the number of self-citations and the number of backward citations (e.g. Fleming, 2001), excluding them from the citation count. We also control for the size of knowledge base which is measured as the sum of the number of patents granted to the firm in the previous 3 years (Yayavaram and Ahuja, 2008). We complete introducing number of technological classes.
We have also data on *number of inventors* and *number of applicants* but we exclude them given their high and statistically significant correlation with *number of patents*. However, it is important to notice that number of patents takes into account also for these two elements. We also introduced several controls at the firm level. We use number of employees for measuring *firm size* in order to control for the scale and scope effects within technological search (Henderson and Cockburn, 1996). Firms that invest more in R&D may generate more innovations. For this reason, we also control for *R&D intensity*, calculated as R&D expenses divided by net sales (Yavavaram and Ahuja, 2008). In Table 1 we provide descriptive statistics and correlations.

<table>
<thead>
<tr>
<th>Insert Table 1 here</th>
</tr>
</thead>
</table>

**Results**

To test hypothesis 1, we used a linear regression model (OLS) with fixed effects to account for time invariant characteristics of firms, thanks to the panel structure of our data. We preliminary test the usual hypothesis related to OLS such as multicollinearity. Moreover, as a robustness check we control for heteroscedasticity and normality of errors clustering results for *firm id*. Table 2 shows results related to our hypotheses.

<table>
<thead>
<tr>
<th>Insert Table 2 here</th>
</tr>
</thead>
</table>

Model 1 indicates support for hypothesis 1. Firms with higher level of centrality are less distant from the optimal structure of decomposability. Eigenvector centrality measures how a firm is related to other central firms. Having a high value of eigenvector means being in the center of industry network, giving access to more knowledge on what are the real interdependencies that exist in the search space.

Model 2 indicates support for hypothesis 2. High values of closeness push the firm far from the optimal degree of decomposability of the knowledge base. This is due to the fact that being too much closer harm the ability of firms to perform distant search that is fundamental to activate change in the knowledge structures. Thus, the knowledge base of a firm with higher values of closeness tend to persist over time and it suffers from an early stage imprinting (Yavavaram and Ahuja, 2008)

Model 3 indicates support for hypothesis 3. Firms holding high number of structural holes toil to receive all the information on couplings produced by the industry network. So they toil to
update their guesses on coupling and this is reflected in their knowledge bases that tend to be distant from the optimal degree of decomposability.

Since a largely debate exists on the role played by structural holes we also check for a curvilinear relationship between structural holes and distance from the optimal degree of decomposability. However, while the linear term remains statistically significant, the quadratic term is not. In other words, while in other contexts structural holes can be also beneficial for firms in accessing new knowledge to perform innovation, when the problem is to establish the optimal fit in the knowledge structures the impossibility of assessing the performance of different knowledge structures caused by structural holes has a negative impact that has no clear curvilinear effects (i.e. the high the number of the structural holes the greater the distance from the optimal degree of decomposability).

For robustness check we also run a regression including all our independent variables (Model 4). Statistical significance stays, confirming our results. Moreover, we run a set of logit regression where the dependent variable is not the distance from the optimum but the fact that the firm’s knowledge base decomposability lays in the interval 0.4-0.6 (where the optimal level should lay). Our results are confirmed.

Discussion and conclusion

We have tried to understand how firms can create nearly decomposable knowledge bases. We have focused our attention on the role played by alliances in the industry wide network as a factor that drives the emergence of a specific structure instead of another one. Our results confirm the idea that alliances favor or limit firms from reaching different degree of decomposability in their knowledge bases, influencing their cognitive representations of the search space. Moreover, alliances play a key role in reducing or increasing the rigidity to change of different knowledge structures, offering information on better couplings to perform. Indeed, network of alliances can encourage changes offering information on the real interdependencies that exist in the search space. More specifically, different networks positions provide access to information that can be used by firms to recombine their knowledge bases and change their coupling pattern. Networks also reduce the costs of acquiring some set of information and knowledge that is necessary for activating changes in the knowledge bases. In this sense networks play a complementary role to internal activities of R&D. These results are particularly interesting because variations in the degree of
decomposability of a knowledge base explain why some firms are better than others in producing innovations. Moreover, these results explain why firms differ in their decision on couplings even when they face the same set of underlying interdependencies.

This study contributes to literature on innovation, underlining the importance of knowledge bases and their linkage with the network of alliance that is very often overlooked. The evolutionary approach to the study of innovation (i.e. Nelson and Winter, 1982; Dosi, 1984) relies on the concept of firm as developed by Cyert and March. In this world, because of bounded rationality, satisficing organizations use standard procedures to perform R&D. “This process leads to a highly path-dependent search for new technologies” (Stuart and Podolny, 1995). In this kind of literature innovation is seen as an incremental process, made of small improvements on existing technologies. Scholars in this area use the concept of «technological paradigm» to stress the idea that technologies develop along paths, determined by «technical attributes, as well as problem solving heuristics and knowledge and skills contained in a paradigm» (Dosi and Orsenigo, 1988). A technological paradigm is defined by the generic task to which it is applied, the technology it exploits, the physical and chemical properties it uses and the technological and economic dimensions and trade-offs it focuses on. In this world innovation means the improvement of the trade-off related to these dimensions (Dosi, 1984).

However, we have demonstrated that the conceptualization of innovation as a recombination of existing elements (Gilfillan, 1935; Schumpeter, 1939; Basalla, 1988; Henderson and Clark, 1990; Hargadon and Sutton, 1997) cannot leave aside the importance for firms of navigating the search space. In order to do this, firms have to develop a guide (i.e. knowledge base) that is influenced by their position in the network alliance. Only considering the role played by external environment and other actors in the industry we can hope to have a clear picture of the innovation process.

As firms learn and become aware of real interdependencies in the search space, they tend to change their knowledge bases, while not all structures make changes feasible. Here we complete the picture suggesting that position in the industry wide network is associated to different structures. More specifically, we have demonstrated that firms with higher centrality tend to assume a halfway degree of decomposability within their knowledge base. Finally, we contribute to literature on networks analyzing the interplay between network position of a given firm and its ability to frame its knowledge base in a certain way.
References


---

### Tables

#### Table 1

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>S.D.</th>
<th>Min</th>
<th>Max</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Distance from optimal degree of decomposability</td>
<td>0.36</td>
<td>0.14</td>
<td>0.00</td>
<td>0.57</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Eigenvector Centrality</td>
<td>0.03</td>
<td>0.08</td>
<td>-0.00</td>
<td>0.68</td>
<td>-0.255</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Closeness</td>
<td>10320.09</td>
<td>6042.12</td>
<td>2190.00</td>
<td>32814.00</td>
<td>0.218</td>
<td>-0.283</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 Structural Holes</td>
<td>0.58</td>
<td>0.38</td>
<td>-0.78</td>
<td>0.98</td>
<td>-0.085</td>
<td>0.301</td>
<td>-0.330</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Firm performance</td>
<td>0.11</td>
<td>0.24</td>
<td>-1.45</td>
<td>1.73</td>
<td>0.112</td>
<td>0.113</td>
<td>-0.056</td>
<td>-0.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Firm size (log)</td>
<td>10.70</td>
<td>1.39</td>
<td>2.20</td>
<td>12.98</td>
<td>-0.292</td>
<td>0.352</td>
<td>-0.429</td>
<td>0.295</td>
<td>0.158</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 R&amp;D Intensity</td>
<td>14.45</td>
<td>42.27</td>
<td>-12.36</td>
<td>353.23</td>
<td>-0.025</td>
<td>-0.027</td>
<td>-0.058</td>
<td>-0.147</td>
<td>-0.071</td>
<td>-0.098</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 Size of knowledge base</td>
<td>8.22</td>
<td>13.74</td>
<td>1.00</td>
<td>125.00</td>
<td>-0.138</td>
<td>0.366</td>
<td>-0.277</td>
<td>0.080</td>
<td>0.026</td>
<td>0.230</td>
<td>-0.040</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>9 Number of technological classes</td>
<td>3.95</td>
<td>2.27</td>
<td>1.00</td>
<td>10.00</td>
<td>-0.589</td>
<td>0.420</td>
<td>-0.315</td>
<td>0.155</td>
<td>-0.016</td>
<td>0.446</td>
<td>0.041</td>
<td>0.345</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Correlations are in parentheses.
### Table 2

**Results of Fixed Effects Regression Analyses for Distance from optimal degree of decomposability**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eigenvector Centrality</strong></td>
<td>-0.747**</td>
<td>-0.771**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.356)</td>
<td>(0.367)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Closeness</strong></td>
<td></td>
<td>4.27e-06**</td>
<td>4.35e-06*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.79e-06)</td>
<td>(2.58e-06)</td>
<td></td>
</tr>
<tr>
<td><strong>Structural Holes</strong></td>
<td></td>
<td></td>
<td>0.055**</td>
<td>0.0758**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.0311)</td>
<td>(0.0313)</td>
</tr>
<tr>
<td><strong>Firm performance</strong></td>
<td>0.0645</td>
<td>0.0551*</td>
<td>0.0964**</td>
<td>0.0736*</td>
</tr>
<tr>
<td></td>
<td>(0.0441)</td>
<td>(0.0346)</td>
<td>(0.0431)</td>
<td>(0.0436)</td>
</tr>
<tr>
<td><strong>Firm size (log)</strong></td>
<td>-0.0276</td>
<td>-0.00232</td>
<td>-0.0321</td>
<td>-0.0176</td>
</tr>
<tr>
<td></td>
<td>(0.0444)</td>
<td>(0.00863)</td>
<td>(0.0442)</td>
<td>(0.0439)</td>
</tr>
<tr>
<td><strong>Size of knowledge base</strong></td>
<td>0.00105</td>
<td>0.000499</td>
<td>0.000772</td>
<td>0.000922</td>
</tr>
<tr>
<td></td>
<td>(0.000680)</td>
<td>(0.000614)</td>
<td>(0.000669)</td>
<td>(0.000677)</td>
</tr>
<tr>
<td><strong>R&amp;D intensity</strong></td>
<td>-0.000992**</td>
<td>-5.73e-05</td>
<td>-0.000335</td>
<td>-0.000630*</td>
</tr>
<tr>
<td></td>
<td>(0.000351)</td>
<td>(0.000175)</td>
<td>(0.000348)</td>
<td>(0.000347)</td>
</tr>
<tr>
<td><strong>Number of technological classes</strong></td>
<td>-0.06523</td>
<td>-0.0288****</td>
<td>-0.00532</td>
<td>-0.00471</td>
</tr>
<tr>
<td></td>
<td>(0.00798)</td>
<td>(0.00417)</td>
<td>(0.00798)</td>
<td>(0.00787)</td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>0.648</td>
<td>0.402***</td>
<td>0.631</td>
<td>0.455</td>
</tr>
<tr>
<td></td>
<td>(0.485)</td>
<td>(0.0991)</td>
<td>(0.485)</td>
<td>(0.483)</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>342</td>
<td>342</td>
<td>342</td>
<td>342</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.082</td>
<td>0.082</td>
<td>0.082</td>
<td>0.114</td>
</tr>
<tr>
<td><strong>Number of id</strong></td>
<td>60</td>
<td>60</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td><strong>Company FE</strong></td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td><strong>Year FE</strong></td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Standard errors in parentheses; all independent variables are lagged by one year

*** p<0.01, ** p<0.05, * p<0.1