

*29th EGOS Colloquium 2013 Montreal, Canada*

**Central positions and performance in the scientific community.  
Evidences from clinical research projects.**

**FEDERICA BRUNETTA**

Department of Business and Management – LUISS Guido Carli University

[fbrunetta@luiss.it](mailto:fbrunetta@luiss.it)

**PAOLO BOCCARDELLI**

Department of Business and Management – LUISS Guido Carli University

[pboccard@luiss.it](mailto:pboccard@luiss.it)

**ANDREA LIPPARINI**

Department of Management – University of Bologna

[andrea.lipparini@unibo.it](mailto:andrea.lipparini@unibo.it)

## Introduction

This research project is part of a wider research plan aimed at addressing some critical issues related to a better understanding of how network structures and dynamics impact on network performance. We observe network structures, their antecedents and subsequent performance. We build on social network arguments and we highlight how certain ego network's structural attributes help firms to exploit opportunities arising from the surrounding environment, improving their performance. We address the following research questions: What are the effects of centrality on performance in cooperative networks? Which structural characteristics are relevant in clinical research networks operating in highly-regulated settings?

Literature is claiming for a deeper understanding of which processes shape the evolution of network structures over time, as the knowledge of networks remains otherwise incomplete (Zaheer and Soda, 2009). We draw on the assumption that the “normative ideal” network structure should be understood according to the context in which the network is embedded, the nature of the actors and the content of the relationships (Ahuja, 2000). Previous studies have focused on network of competitors; we observe collaborative networks generated by the necessity to respond to meeting legal or regulatory requirements, shaped by institutional forces. Authors have explored the role of actor centrality in competitive networks but, unless some relevant conceptual contributions (Dhanaraj and Parkhe, 2006), the role of centrality in cooperative networks remains largely unexplored. In an attempt to explain network evolution, we observe how past presence of structural holes impacts on the current centrality of an actor we examine the effects of centrality on performance in highly-regulated and cooperative contests.

Our study is set in the context of clinical research projects in the pharmaceutical industry, a setting characterized by institutional forces that significantly constrain actors and is subject to increasing specialization and complexity. We focus on clinical trials that can be conducted locally or internationally, in a single site or across multiple sites. Accordingly, they may be focused on a single institutional context or they may span multiple institutional environments. When trials are conducted across multiple sites, they are connected by the same protocol, and the network of sites can be considered a research network, or a bundle of “innovating firms working together” (DeBresson & Amesse, 1991). Research networks in this industry are likely to provide timely access to otherwise unavailable knowledge and resources, stimulate internal expertise, and

contribute to learning capabilities (Powell, Koput & Smith-Doerr, 1996). Such networks are a consequence of the increasing technological and market uncertainties witnessed by pharmaceutical companies over the last decades, with the subsequent need to leverage upon complementary know-how and other resources.

## **Theory and Hypotheses**

This study focuses on research networks, consisting of “innovating firms working together” (DeBresson and Amesse, 1991). In contrast to other studies that have focused on competitive dynamics, we investigate networks characterized by co-operative dynamics between actors. Within these networks, co-ordination among nodes of their competences and assets serves to enhance the contribution of each node (Powell et al, 1996).

Studies have demonstrated that the position of firms in interorganizational networks influences their behavior and outcomes (Powell et al., 1996). External collaboration is complementary to internal capabilities as it facilitates exploiting and building upon existing knowledge. Networks offer two substantial benefits: sharing resources, knowledge and skills, and access to knowledge spillovers. The former enables the transfer of know-how and physical assets; the latter facilitates the transfer of information (Ahuja, 2000).

The appropriate network structure that allows catching these benefits may depend on a variety of moderating factors. A contingent view on the benefits of network structure is largely being argued (Uzzi, 1997; Ahuja, 2000; Obstfeld, 2005). We identified a contingency factor that is often discussed, but rarely investigated, in the debate about the relationship between network structure and performance (Powell et al, 2005). We focus on collaborative relationships generated by the *necessity* (Oliver, 1990) to respond to meeting legal or regulatory requirements as institutional forces rationalize the environment in which the network is embedded (Uzzi, 1997). Institutions are collective and regulatory complexes that regulate the actions of organizations. Regulation happens through the enforcement of laws, rules and norms that constitute constraints (DiMaggio and Powell, 1983; Child and Tsai, 2005) affecting the behavior of the organization, its routines and its relationships with other firms.

Networks stimulate internal expertise and learning capabilities (Powell et al, 1996). Resources and knowledge sharing account for positive impacts on performance: network acts as a channel through which partners share the knowledge and experience from their interactions with network partners (Gulati and Gargiulo, 1999).

In this study, successful scientific performance is measured by the citations that the principal investigator receives following the beginning of the project he is in charge of. Being the principal investigator the coordinator of the whole study, he can be seen as the *knowledge and information gatekeeper*, monitoring the environment, translating and collecting the information for the whole research group (Cohen and Levinthal, 1990). The investigator's (actor's) centrality across the network may impact on its capability of being an information gateway by disseminating and receiving information and knowledge. The central actor can be strategically important, playing a pivotal role: enhancing knowledge mobility, ensuring equitable distribution of value, and fostering trust among actors and finally promoting network stability (Dhanaraj and Parkhe, 2006).

Literature is claiming for a deeper analysis of network evolution (Powell et al., 2005; Zaheer and Soda, 2009) at the dyadic and network level. In this light, authors have explored the role of actor centrality in a longitudinal perspective looking at centrality as an antecedent of specific network structure characteristics (Zaheer and Soda, 2009). Nevertheless, the investigation of the antecedents of actor centrality in network dynamics remains largely unexplored.

In cooperative networks partners are encouraged to share resources and knowledge (Gargiulo and Benassi, 2000) in order to meet individual and network outcomes. Spanning structural holes is strategically important for the circulation of information: the more the holes spanned the richer the information benefits, as an access to a broader information screen enhances the chance to identify opportunities and guarantees a timely gathering of information (Burt, 1992; Powell et al., 1996). Actors who span structural holes can have a speedy access to diverse information in different parts of the network that might be difficult to reach otherwise.

In a context of cooperation in which coordination of action is fundamental to the innovation outcome, actor spanning a structural hole may act as *tertium iungens* promoting the creation and facilitation of ties among alters in order to foster the innovation activity and maintaining a coordination rather than mediation role (Obstfeld, 2005).

We argue that this may lead the actor to increase its centrality, and hypothesize:

*H1 in cooperative settings egos spanning structural holes will produce more central positions. Specifically, the higher the past structural holes spanned by an actor the higher is its centrality.*

Since the idea of centrality has been proposed by Bavelas (1948), central positions have always been associated to higher influence and privileges (Freeman, 1979), prominence and visibility (Wasserman and Galaskiewicz, 1994), power (Brass and Burkhardt, 1993) or prestige (Bonacich, 1972). The key importance of centrality is related to superior resource and information advantages: the power arising from central position differs from the power obtained by the actors' attributes and confers substantial advantages (Rowley, 1997). High centrality is associated to a higher volume and rapidity of flows of information and resources, opportunities.

Nevertheless, environmental constraints, strategic intent and the actions carried-out by network actors may impact on the benefits arising from network position. Centrality may have two negative consequences: first, a central actor is related to a large number of ties and this can overturn in a highly dependence upon its network; second, each tie represents for the actor a source of information and resources but also a weak point through which those could drain. As centrality increases, projects are likely to experience augmented coordination costs (Gulati and Singh, 1996; Gargiulo and Benassi, 2000), and cognitive limits (Hansen et al., 2005) which contribute to decreasing the benefits of centrality. A position of visibility into the collaboration network may translate into a constraint on the ego network's ability to absorb new information and ideas or to respond as flexibly as companies with few direct ties.

These negative sides of centrality are usually outweighed or balanced by the benefits that the central actor gains for occupying privileged positions. This balance holds for situations in which the actors are not only related by cooperative relations but also by competitive dynamics (Rowley, 1997, Dhanaraj and Parkhe, 2006): as noted by Burt (1992), the position alone does not create the benefit, as the advantages are determined by the entrepreneurial behaviour of the actor.

Accordingly we hypothesize:

*H2 In cooperative networks, the higher the centrality of an actor, the lower its performance.*

Literature claims a deeper analysis of the contingent effects of the institutional constraints on the relationship between network structure and performance (Powell et al, 2005). We define institutional embeddedness as the interconnections between organizations and the institutional environment (DiMaggio and Powell, 1983). We analyze the mechanism through which institutions and rules shape the behaviours, actions and relations. Institutional forces rationalizing the environment in which the network is embedded may leverage network relations or break social ties and the consequent benefits of embeddedness (Uzzi, 1997).

In an highly regulated research context, in which the research process is strictly monitored and procedural conformity arises, the development of a different sets of perspectives, skills and resources (Burt, 1992) and information boosting the scientific production might be limited, and the coordination role carried out by the central actor may account for mechanisms associated to coordination costs. The research process is less likely to have a clear scheduling as the outcome emerges as a result of changing conditions and different applicable regulations: adjusting mechanisms are associated with changing coordination mechanisms (Thomassen and Lorenzen, 2001). When the environment is highly regulated, as procedural nuances have to be followed, firms tend to conform (Scott, 1987), driven by procedural conformity that forces organizations in the way they perform specific tasks or procedures following specific requirements that are imposed by institutions. We argue that when actors are embedded in different institutional environments, they are offered the possibility to catch knowledge and opportunities residing in other parts of the network. Diverse institutional spheres can represent windows in which frameworks tend to be less constraining for the actors.

We hypothesize:

*H3: In cooperative networks, the effect of centrality on performance will be moderated by the presence of regulation constraints.*

### **Empirical setting**

We test our hypotheses in a cooperative network made of 114 research projects carried out in the pharmaceutical industry. Over the last few decades, this setting evolved towards high

specialization and high systemic complexity. Moreover, in this specific context governmental actors and strict institutional regulations significantly constrain actors' behavior. Specifically, we look at the clinical trial research networks. Clinical trials can be conducted locally or internationally, in a single site or across multiple sites. Accordingly, they may be focused on a single institutional context or they may span multiple institutional environments. When trials are conducted across multiple sites, they are connected by the same protocol, and the network of sites can be considered a research network. Multi-center research affects important aspects of trial performance, as well as the extent to which the study can achieve its aims, by obtaining the maximum amount of data to be collected (Mattesich, Murray-Close & Monsey, 2001; Greene, Hart & Wagner., 2005)

Drug research proceeds through different phases (Smith & O'Donnell 2006), strictly monitored by regulations. In the pre-clinical phase, the potential lead compounds are synthesized and tested on cultured cells and animals to assess toxicity or efficacy. Once this phase is completed, the lead compounds are reduced to a few useful candidate drugs and advanced to the clinical development stage. Clinical testing proceeds through three phases. In phase I, the compound is tested on a small number of healthy volunteers to establish safe dosages and gather information on possible side and metabolic. Phase II consists of randomized, double blind studies to ensure objectivity, evidence on safety, and preliminary data on efficacy. It involves a larger group of volunteers affected by the target disease. Phase III consists of a large-scale study, designed to collect effective information concerning the treatment's safety and effectiveness and involving up to thousands of volunteers. Finally, once the evidence on safety and efficacy is gathered, the developers submit the results to regulatory authorities for marketing approval.

Networks of clinical research centers are composed of centers working on common projects on a multi-center basis in order to enroll a sufficient number of patients and share scientific knowledge. Following Huckman and Zinner (2007), each of these investigative sites can be considered as a 'factory', where the end-product is a fully evaluated study subject. The protocol and the contract with the sponsor bind the action of sites (Huckman & Zinner, 2007). The former dictates the guidelines of the product (trial), that is identical across all sites in a given trial; the latter defines the product quantities (enrollment goal) and the price (budget provisions for each subject), as well as the expected delivery (enrollment period). Despite the constraints dictated by the agreement, each site maintains the capability of formulating operational and

management decisions concerning the access to the material (study subjects), allocation of the activities among the team (principal investigator, sub-investigators and study coordinators), and the operational execution (enrolling and processing subjects).

The pharmaceutical industry is also characterized by a strong effort in publishing in the scientific literature (Powell et al., 1996), aiming at commercializing basis of scientific discovery and emphasizing the relation between knowledge and innovation output (Gittelman & Kogut, 2003). Each trial is coordinated by a principal investigator, responsible of the study outcome for all the sites involved, as well as the administrative and bureaucratic activities that have to be carried on in order to conduct the trial in respect of the regulation and publicity of the results. Being the principal investigator the coordinator of the whole study, he can be seen as the knowledge and information gatekeeper, monitoring the environment, translating and collecting the information for the whole research group (Cohen & Levinthal, 1990).

### **Data, variables and measures**

This research uses data gathered from national and international databases on clinical trials, including single and multi-center clinical trials, both national and international, so as to construct a full network of actors engaged in research in the same field and time period, 466 projects were identified on Clinicaltrials.gov, according to the specific clinical fields of viral diseases (specifically, Hepatitis-related trials) performed from January 2000 to 1 June 2009—overall, these projects involved 2090 centers. We then analyzed data on the 114 projects sponsored by private industry in which the recruitment of patients has already been closed, therefore on project characterized by a “closed recruitment status” (and only such trials), allowing for comparison among different levels of study performance.

The analysis focused on projects carried out in centers located in 8 selected countries (Germany, France, Italy, Netherlands, UK, US, Canada and Australia). These countries were selected as the main hosting countries for clinical trials (Thiers, 2006) and major pharmaceutical markets.

Clinical trials projects respect the underlining assumptions that have to be made in order to shift the network level of analysis from individual to organization, as noted by Zaheer and Soda (2009). The content of the ties between projects can be thought of in terms of resources (i.e.,

basis of enrolled patient), and competences, (i.e., procedural competences). First, clinical research projects respect the assumption of network composition – a single link connects one project to the other (i.e., an investigative site) – while coordination processes among sites working on a project are tight and strictly regulated, and therefore a site working within another project is meant to affect the members of the project as a whole. At the same time, as these two dimensions are related, clinical research projects respect the assumption of contagion: if two projects share an investigative site, and one of these two projects shares another investigative site with a third project, then the third project is likely to be affected by practices that are being used in the project with which it is not directly linked. Nevertheless, this process can be mediated by the interfaces and coordination processes within projects, also emphasized by a procedural conformity (Scott, 1987) arising from the surrounding institutional environment. Finally, these projects respect the assumption of causality – i.e. the structural content of the network translates into the performance of each project – along two dimensions: network allows the identification of resource-sharing opportunities and learning processes both in the transfer of knowledge on efficient procedures and institutional opportunities, and knowledge on research-related topics.

Data on the 114 projects have been gathered from a variety of sources. Data on trials have been collected on Clinicaltrials.gov and its archive (Clinicaltrials.gov archive). Publications data were found on Pubmed.gov. Data on the regulations in act in the eight selected countries have been collected through the analysis of single codes, guidance and regulations. For the five EU countries, records and studies from EMEA have been used as a comparison. Project attributes have been collected on Clinicaltrials.gov (number of location countries, number of sites), while we addressed to the sponsors websites to gather sponsor attributes. Similarly, attributes for each clinical trial center have been collected from the websites of every facility, except for the number of beds, that in some cases is collected nationally on specific databases as in Germany (Kliniken.de and Krankenhaus.net), France (Hopital.fr), Italy (Ministry website) and the US (American Hospital Directory and Hospitaldata.com). Finally, data on the disease incidence have been gathered through specialized websites such as the ‘WHO ICTRP’ or disease-specific association of practitioners.

Network data have been collected through the Clinicaltrials.gov database and integrated with data from the ‘WHO ICTRP’ and the ‘IFPMA’ clinical trials portal; the records for each project contain a list of the sites where the trial is being conducted and the name of the principal

investigator which is in charge of the study. The location is easily identifiable in case of full records; nevertheless, since the specific identity of the medical center in which the site is located is sometimes reported partially, every site has been identified by a research that has cross-referenced facility and a confirmatory research within the website of the facility of current areas of trial investigation.

### **Analysis and econometric approach**

We first used a 2SLS model to test our first two hypotheses. We used a robust variance estimator to control for the effects of correlation between errors across equations due to endogeneity between network structure and performance. We run several tests to check the consistency of our modeling approach. In particular, we run a Wu-Hausman F-Test and the Durbin-Wu-Hausman  $\chi^2$  test: these tests serve to verify or reject the null hypothesis that states that an ordinary least squares (OLS) estimator of the equation would yield consistent estimates, and thus endogeneity among the regressors would not have deleterious effects on OLS estimates.

The third hypothesis then used a regression model to estimate the moderating effect of regulation constraints on the effect of centrality on performance.

#### ***Two-stage Least Square Analysis (2SLS): First-stage Variables***

All network data and measures were analyzed and calculated through the UCINET 6.0 software (Borgatti, Everett and Freeman, 2002). Network measures for each of the 114 analyzed projects have been calculated over the full network of 466 projects. Network data were first collected in matrices “actor x project” identifying the ties between the sites and the projects. Through the UCINET 6.0, the main matrix was then transformed (cross-products, co-occurrence method) in squared, one mode affiliation matrices “project x project”, in order to create primary affiliations data and to calculate specific network indicators referring to the single projects. This network can be considered “isomorphic” with research on individuals as network nodes, as the composition, contagion, and causality assumptions are made (Zaheer and Soda, 2009, p. 12).

### ***Endogenous Variable: Current Centrality***

To measure our endogenous variable, we considered the measure of Closeness Centrality. In networks of cooperative research projects, a process of efficient communication, involving fewer message transmissions, shorter times and lower costs to access independently all other members of the network (Freeman, 1979) can be fundamental in order to gain information, resource advantages and foster the innovation performance. This advantage is identified in literature by the idea of closeness centrality (Freeman, 1979) that measures the ability of an actor to reach other actors through a minimum number of intermediary positions (Brass, 1984).

Symmetrically, the central actor spreads information in the network through fewer intermediaries than those actors that occupy a peripheral position: being "close" to all the others, the central actor can quickly access the network (Freeman, 1979), enhancing the mechanism of cooperation. We therefore use closeness centrality as an indicator of centrality. Closeness centrality is based upon "the frequency with which a point falls between pairs of other points on the shortest or geodesic paths connecting them" (Freeman, 1979). 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. We computed this measures using UCINET VI (Borgatti et al., 2002).

### ***Instrumental Variable***

*Past structural holes.* Our input dataset consisted of 1280 clinical trial centers and 114 total projects. Following the work of Zaheer and Soda (2009) we have constructed our past network. We've first composed 7 matrices corresponding to each moving window comprising two years (starting from the window 2001-2002 and ending with the window 2007-2008). Specifically, we used a time window of two years which corresponds to the longest past we could obtain using 2003 as a cutoff that could allow us to monitor short-term effects. Following this procedure (suggested by Zaheer and Soda, 2009) we were able to split the network with regard to the past and the current time and thus moving the two-year windows across multiple years.

We constructed an individual vector for each project (1280 centers actors  $\times$  1 project), excluding the project of the first window (2001-2002). We've then created multiple matrices, composed by

the past network window to which we added, repeatedly, the vectors, each of them representing a focal project. Finally, using UCINET version 6.0, these matrices were transformed (co-occurrence method) into squared, one-mode affiliation matrices “project per project”. On these co-membership matrices we calculated the specific ego network indicators.

Specifically, we calculated for each of the focal project the measure of *Past structural holes* given by the measure of *efficiency* of the focal project in its past network.

### ***Control variables***

***Clinical center-related variables.*** We use the size of clinical centers as factor influencing trial quality and network performance. Starting from the matrix actor per project, we identified a list of participating centers in the eight analyzed countries. We have used a Log transformation of the *Cumulative Size of clinical centers participating in the project* referring to the cumulative size of the health facilities involved in the trial. Based on prior literature (Dranove, 1988, Goodstein, Gautam & Boeker, 1994; Goodstein, Boeker & Stephan, 1996; Chadwick, Hunter & Walston, 2004), this is measured through the number of licensed hospital beds, a common size measure for hospitals and healthcare facilities and one that allows a cross-country comparison.

***Project-related indicators.*** We control for the *Number of Location Countries* in which the trial is taking place, which is an important measure of the geographic spread of the trial as countries that may represent important sources of information and knowledge. Similarly, we control for the *Number of centers participating in the trial*, in order to control for the number of clinical trial centers who contribute to the same project, coordinated by the principal investigator.

We take into account the size of the trial, measured through the *Enrollment target*, that accounts for the number of patients to enroll into the trial. Finally, we control for the trial Phase (*Phase I, Phase II, Phase III*) as the latest trial phases are usually associated with an higher number of publications and citations as they represent a more advanced research status.

***Sponsor-related indicator*** We use the Log transformation of the *Years of operations of the firm sponsoring the project* referring to the experience in clinical trials. Firm age is used as a control variable to capture the effects of firms’ prior work experience and learning advantages on outcomes. Firm age is measured as the cumulative number of years of operation at the beginning of the trial.

***Principal investigator related indicators.*** We control for the tenure status of the principal investigator. Specifically, this measure takes into account whether the principal investigator is an: *Associate Professor* (the principal investigator has reached the level of Associate Professor), *Full Professor* (the principal investigator has reached the level of Full Professor), *Head or Director* (the principal investigator has reached the position of head of an academic department or is the director of a specific unit within a pharmaceutical firm). We also control for the number of publications issued before the beginning of the trial (*Previous publications*).

### **Two-stage Least Square Analysis (2SLS): Second-stage Variables**

On the second stage we aim to estimate the impact of *current centrality* (our endogenous variable) on performance.

*Dependent variable: number of citations.* The scientific performance and is measured through the number of citations. Bibliometric measures are widely used to evaluate the impact of research outcomes. We use the number of citations (Judge et al, 2007) accrued for each article published by the principal investigator in the two years following the date of beginning of the trial.

### **Control variables**

The following variables are relevant to control for the role of different factors on the scientific performance of the principal investigator:

***Project-related indicators.*** In order to control for the complexity of the trial and the coordination activity needed from the principal investigator. We control for the *Number of Location Countries* in which the trial is taking place, for the *Number of centers participating in the trial*, coordinated by the principal investigator and for the size of the trial, measured through the *Enrollment target*. We also control for the trial Phase (*Phase I, Phase II, Phase III*) as the latest trial phases are usually associated with higher complexity and number of publications.

***Principal investigator related indicators.*** We control for the number of publications issued before the beginning of the trial (*Previous publications*), as this variable may impact on the number of citation accrued by the principal investigator.

### **Regression Analysis**

In order to estimate the moderating effect of regulation constraints on the effect of centrality on performance we used a regression model, using *number of citations* as our dependent variable and *closeness centrality*, as the independent variable.

The moderator is the measure of *Regulation Constraints*. Our empirical work involved conducting an extensive analysis of the regulatory frameworks across the selected countries, as a basis for creating an independent variable to measure the institutional constraints for each country. We have considered the type and extent of activities firms were not permitted to engage in and the administrative burdens imposed as a measure of the constraint imposed by the regulatory frameworks (Barth et al., 1997). We identified four regulation components affecting the recruitment process as factors of constraint within the institutional frameworks. Specifically, we consider:

*Document Numbers*. Differences in the administrative burdens exist between different analyzed countries. Some, such as the US or Australia (where clinical trials are evaluated via the CTX authorization process), have a lower document load that, for example, Italy and the Netherlands, where the clinical trial application (CTA) process requires more paperwork.

*Response times*. Responses by Competent Authority (CA) are compulsory before starting the enrollment process –differences exist in the time of response across the studied countries, which may delay the process.

*Advertising trials*. This can be very important for recruitment participants, especially where payment is allowed. Some countries allow listings of trials and recruitment sites (aside from public registries) where volunteers and patients can gather information about and apply to join the trial. In others, trials are only publicized through the national registries and enrollment is proposed directly to participants by physicians.

*Payments to participants*: Some countries allow participants to be paid; which represents an incentive to participants, and thus impacts on the speed of recruitment.

To create a regulation index, we first standardized the components by calculating the distance from the mean  $X_i - \bar{X}$  divided by the standard deviation (the dummy variables "Advertising trials" and "Payments to participants" were coded 0 for "Yes" and 1 for "No"), in

order to reach a comparable scaling among items. We have weighted each component for every project by the percentage of centers embedded in the specific country context participating in the project (CEz) of the total number of the participating actors (CE). We then proceeded with an exploratory factor analysis (EFA) to understand the underlying factor structure. We have used the Kaiser-Meyer-Olking test to evaluate the factorability of the correlation matrix, that appears to be appropriate for factoring as the KMO measures are 0.70 and higher. In order to verify the number of factors to be extracted we double checked using the Kaiser criterion, that suggested to retain a single factor with an eigenvalue higher than 1 (3.07) and then run a parallel analysis, that proved that the four items consistently load on only one factor. Finally, we examined factor loadings and retained all the initial items.

Selected control variables are the same used in the second stage of the Two-stage Least Square Analysis (2SLS) model.

## Results

In order to verify the appropriateness of treating structural holes as an endogenous variable, and to reject the null hypothesis that current structural holes are exogenous to performance, we have run the Wu-Hausman F-test [2.49, d.f. (1,52);  $p = .12$ ], and the Durbin-Wu-Hausman  $\chi^2$  test [2.74 (1);  $p = .09$ ]. Tests confirm the appropriateness of using a 2SLS specification to address the issue of endogeneity.

We have first used a 2SLS model. We used the Durbin-Watson test to check past structures for autocorrelation with current structures, and the test confirmed that there was no autocorrelation. We corrected for the presence of heteroskedasticity by using the Huber-White sandwich estimator of variance in Stata (Huber, 1967; White, 1980).

Descriptive statistics are reported in table I. We report the results of the 2SLS analysis in table II and III of both first and second stages. The second stage of the 2SLS tests the effects of current structural holes on project performance. We explain the results of the model below.

-----  
INSERT TABLE I and II ABOUT HERE  
-----

The model shows the effect of the instrumental variable Past structural holes. The results confirm our first hypothesis as past structural holes impact positively on current centrality, in the 2SLS model  $\beta = .10$  significant at  $p < .1$ . Therefore, we find evidence that in cooperative settings egos spanning structural holes will produce more central positions. Specifically, the higher the past structural holes spanned by an actor the higher is its centrality. We have interpreted these results on the basis of the peculiar context of cooperation, where the coordination of action is fundamental to the innovation outcome and actors spanning structural holes may act as *tertium iungens* promoting the creation and facilitation of ties among alters in order to foster the innovation activity and maintaining a coordination rather than mediation role (Obstfeld, 2005).

-----  
INSERT TABLE III ABOUT HERE  
-----

Table III show the result of the second stage regression model on project performance. We aimed at testing that in cooperative networks, the higher the centrality of an actor, the lower its performance. We argued that, as centrality increases, projects are likely to experience augmented coordination costs (Gulati and Singh, 1996; Gargiulo and Benassi, 2000), and cognitive limits (Hansen et al., 2005) which contribute to decreasing the benefits of centrality. A position of visibility into the collaboration network may translate into a constraint on the ego network's ability to absorb new information and ideas or to respond as flexibly as companies with few direct ties. Results indicate a negative impact of current centrality (predicted value) on performance ( $\beta = -49.80$ ). Our second hypothesis is therefore confirmed, although significance is low.

Our final hypothesis aimed at testing the moderating effect of regulation constraints on the negative effect of centrality, in the context of cooperative networks. Institutional forces rationalize the environment in which the network is embedded (Uzzi, 1997) so we focused on the contingent effects of the institutional constraints on the relationship between network structure and performance (Powell et al, 2005).

-----  
INSERT TABLE IV ABOUT HERE  
-----

The results of the regression model are presented in Table IV (significant at  $p < .0000$ ). The results confirm our hypothesis. In Model 3, the full model, the main effect of Centrality is negative and significant, at  $p < .1$ . The main effect of the Regulation Constraints is positive and significant, at  $p < .1$ . The interaction of Centrality and Regulation Constraints is positive and significant at  $p < .1$ . The results show that Regulation Constraints moderate the negative effects of centrality. This confirms our H3.

### **Discussion and conclusion**

Our analysis, rooted in the social network literature, further contributes to the debate on ideal social structures. We provide evidence that in collaborative networks, an actor's centrality is likely to increase according to its past structural holes. Moreover, we observe that an actor's centrality has a negative effect on performance, and that this effect is moderated by the regulation constraints.

These results have implications at the firm, network, and environmental levels. The effects of centrality on performance prompt us to suggest a deliberate shaping of network structures by prominent actors aimed at either improving their central position into a network *and* increasing the benefits arising from the total number of connections existing at the network level. The first action asks for both explorative and exploitative capabilities. The former legitimate a company as innovative, therefore potentially trustable as a competent and reliable partner within a knowledge network; the latter demonstrate to the environment that the actor is able to leverage upon its current knowledge base, valuing his past abilities to reach a desired end. Both these capabilities are conducive to more central positions into a collaborative network, providing stimuli to other potential partners to effectively engage in valuable projects. The second action asks for relational capabilities and brokering attitudes, in order to increase density and create a bigger market for ideas. It is important to clarify that no direct link exists between an increase in the number of ties and innovation at the network level. Nevertheless, variety is usefully alimented by differential contributions, and the quality of the actors involved do the rest. This means, at the firm level, relying upon resources able to scan the environment, to seize it, and to reconfigure it in function of the desired end (Teece, 2007). Many companies in the pharmaceutical and in the airline industry

are staffing their structures with alliance managers who usefully complement the set of capabilities conducive to high performance.

The ability to cope with external constraints is a central managerial implication from our research. When the environment is challenging and with fewer rooms to maneuver, the inability to escape from centrality may be a source of either idea and action problems therefore leading to rigidity at the project level. In these cases, ties can bind and might be disentangled to cope with a new highly regulated setting or with an increase in the limits posed by an existing regulatory framework.

Our results suggest the need for the creation or the affirmation of a behavioral inclination at creating or facilitating ties among firms instead of keeping them far apart (as the *tertius gaudens* approach would suggest). More precisely and importantly, what seems to matter is the *tertius iungens* strategic orientation (Obstfeld, 2005) as the means for moderating not only the magnitude of density in a collaboration network, but also for spanning boundaries and realigning the network with the environment in which it operates. Evidences suggest, one more time, how critical is the role of a coordinative agent, especially for companies subjected to regulatory limits to their action, in avoiding the fragmentation of ideas, processes, all potentially valuable in the search of fit with the environment.

These findings are part of a research plan aimed at addressing some critical and promising issues. We are collecting longitudinal data for a better understanding of how network structures and dynamics would impact on network performance and in order to identify patterns of homogenization among network actors that may arise in institutionalized settings. On the operative side, we plan a fine-tuning of the control variables with the inclusion of new control sets, such as other indicators of the investigator capabilities and sponsor-related indicators and we plan to enlarge our sample including projects conducted in emerging countries (e.g. India, China, Mexico) where clinical trials are increasingly taking place.

The route towards the enhancement of this firm's ability to scan, seize and reconfigure the environment to moderate regulatory constraints, might require a change of the traditional managerial mindsets. The goal here is the development of a greater "mutual interest" (Williamson, 1985), as well as the development of a sense of "common destiny". Leadership is a key ingredient in achieving this end (Fleming & Waguespack, 2007), as well as the use of interpersonal coordination mechanisms (March & Simon, 1958; Galbraith, 1973). A clear

communication of goals, and an appropriate set of incentives to encourage actors in the knowledge network to share valuable knowledge with each other, will encourage them to set up idiosyncratic knowledge-sharing routines, to further facilitate the learning of specified and agreed-on information and know-how between them (Dyer & Nobeoka, 2000).

## References

- Ahuja G. 2000. Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study, *Administrative Science Quarterly*, Vol. 45: 425-455
- Barth, J. R., Nolle, D. E., & Rice, T. N. (1997). Commercial Banking Structure, Regulation, and Performance: An International Comparison. *Managerial Finance*, 23(11), 1-39.
- Bavelas A. 1948. A Mathematical model for group structures. *Human Organization* 7(3): 16-30
- Bonacich P. 1972. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology* 2: 113-120
- Borgatti SP, Everett MG, Freeman LC. 2002. *Ucinet for Windows: Software for Social Network Analysis*. Analytic Technologies: Harvard
- Brass, D. J. (1984). Being in the Right Place: A Structural Analysis of Individual Influence in an Organization. *Administrative Science Quarterly*, 29(4), 518-539.
- Brass, D., & Burkhardt, M. (1993). Potential power and power use: An investigation of structure and behavior. *Academy of Management Journal*, 36(3), 441-470.
- Burt R. 1992. *Structural Holes: The Social Structure of Competition* Harvard University Press: Cambridge, MA
- Chadwick, C., Hunter, L., & Walston, S. (2004). Effects of downsizing practices on the performance of hospitals. *Strategic Management Journal*, 25(5), 405-427.
- Child, J., & Tsai, T. (2005). The Dynamic Between Firms' Environmental Strategies and Institutional Constraints in Emerging Economies: Evidence from China and Taiwan\*. *Journal of Management Studies*, 42(1), 95-125.
- Cohen, W., & Levinthal, D. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1), 128-152.

- DeBresson, C., & Amesse, F. (1991). Networks of Innovators - a Review and Introduction to the Issue. *Research Policy*, 20(5), 363-379.
- Dhanaraj, C., & Parkhe, A. (2006). Orchestrating innovation networks. *Academy of Management Review*, 31(3), 659-669.
- DiMaggio, P., & Powell, W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48, 147-160.
- Dranove, D. (1998). Economies of scale in non-revenue producing cost centers: implications for hospital mergers. *Journal of Health Economics*, 17(1), 69-83.
- Dyer, J., & Nobeoka, K. (2000). Creating and managing a high-performance knowledge-sharing network: the Toyota case. *Strategic Management Journal*, 21(3), 345-367.
- Fleming, L., & Waguespack, D. (2007). Brokerage, boundary spanning, and leadership in open innovation communities. *Organization Science*, 18, 165-180.
- Freeman, L. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1(3), 215-239.
- Galbraith, J. (1973). *Designing Complex Organizations*. Reading, MA: Addison-Wesley.
- Gargiulo, M., & Benassi, M. (2000). Trapped in Your Own Net? Network Cohesion, Structural Holes, and the Adaptation of Social Capital. *Organization Science*, 11(2), 183-196.
- Gittelman, M., & Kogut, B. (2003). Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science*, 49(4), 366-382.
- Goodstein, J., Gautam, K., & Boeker, W. (1994). The Effects of Board Size and Diversity on Strategic Change. *Strategic Management Journal*, 15(3), 241-250.
- Greene, S., Hart, G., & Wagner, E. (2005). Measuring and improving performance in multicenter research consortia. *JNCI Cancer Spectrum*, 2005(35), 26.

- Gulati R, Singh H. 1996. The Architecture of cooperation: managing coordination costs and appropriation concerns in strategic alliances, *Administrative Science Quarterly*, Vol. 43: 781-814
- Gulati, R., & Gargiulo, M. (1999). Where Do Interorganizational Networks Come From? 1. *American Journal of Sociology*, 104(5), 1398-1438.
- Hanneman, R., & Riddle, M. (2005). Introduction to social network methods. Riverside, CA: published in digital form at <http://faculty.ucr.edu/~hanneman/>
- Hansen, M., Mors, M., & Lovas, B. (2005). Knowledge sharing in organizations: Multiple networks, multiple phases. *Academy of management Journal*, 48(5), 776-793.
- Huber, P. J. (1967). "The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. I, pp. 221–33.
- Huckman, R., & Zinner, D. (2007). Does focus improve operational performance? Lessons from the management of clinical trials. *Strategic Management Journal*, 29(2), 173-193.
- Judge, T., Cable, D., Colbert, A., & Rynes, S. (2007). What causes a management article to be cited-article, author, or journal? *Academy of Management Journal*, 50(3), 491.
- March, J. G., & Simon, H. (1958). *Organizations*. New York: Wiley.
- Mattessich, P. W., Murray-Close, M., & Monsey, B. R. (2001). *Collaboration: What makes it work? A Review of Research Literature on Factors Influencing Successful Collaboration*. St. Paul: Wilder Publishing Center.
- Obstfeld, D. (2005). Social networks, the tertius iungens orientation, and involvement in innovation. *Administrative Science Quarterly*, 50(1), 100-130.
- Oliver, C. (1990). Determinants of Interorganizational Relationships: Integration and Future Directions. *Academy of Management Review*, 15(2), 241-265.

- Perks, H., & Easton, G. (2000). Strategic alliances: Partner as customer. *Industrial Marketing Management*, 29(4), 327-338.
- Powell, W. W., & DiMaggio, P. J. (1991). *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.
- Powell, W. W., & Smith-Doerr, L. (2005). Networks and economic life. In N. Smelser & R. Swedberg (Eds.), *Handbook of Economic sociology*. Princeton: Princeton University Press.
- Powell, W. W., Koput, K. W., & Smith-Doerr, L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly*, 41(1), 116-145.
- Powell, W., White, D., Koput, K., & Owen-Smith, J. (2005). Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences 1. *American Journal of Sociology*, 110(4), 1132-1205.
- Rowley, T. (1997). Moving beyond dyadic ties: A network theory of stakeholder influences. *Academy of Management Review*, 22(4), 887-910.
- Scott WR. 1987. *Organizations: Rational, Natural, and Open Systems*. Prentice-Hall: Englewood Cliffs
- Smith, C., & O'Donnell, J. (2006). *The Process of New Drug Discovery and Development*. New York: Informa Healthcare.
- Teece, D. J. (2007). Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, 28(13), 1319-1350.
- Thiers, F. A. (2006). *The globalization of clinical drug development*. Boston: Massachusetts Institute of Technology.
- Thomassen M, Lorenzen M. 2001. The dynamic costs of coordination and specialization, *DRUID Nelson and Winter Conference, Aalborg, June*:

- Uzzi B. 1997. Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness., *Administrative Science Quarterly*, Vol. 42:
- Wasserman S, Galaskiewicz J. 1994. *Advances in social network analysis: research in the social and behavioral sciences*. SAGE: Thousand Oaks
- White, H. (1980). "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48, 817–38.
- Williamson, O. E. (1985). *The economic institutions of capitalism: Firms, markets, relational contracting*. New York, London: Free Press.
- Zaheer A, Soda G. 2009. Network Evolution: The Origins of Structural Holes, *Administrative Science Quarterly*, Vol. 54: 1-31

## Tables

**TABLE I**  
**Descriptive statistics and correlations**

Variables	Mean	S.D.	1	2	3
1 Number of centers participating in the trial	15.48	28.15			
2 Number of Location Countries	2.86	3.98	.775***		
3 Phase	.00	.97	-.147*	-.160*	
4 Enrollment target	655.30	2228.47	.254***	.161*	-.145*
5 Previous publications	25.34	40.68	.038	-.022	-.076
6 Tenure Status	.74	.66	.083	.135	.017
7 Past structural holes	.72	.22	.086	-.062	.096
8 Cumulative Size of clinical centers participating in the project	3.09	.75	.659***	.534***	-.145*
9 Years of operations of the firm sponsoring the project	1.90	.44	-.074	-.053	-.382***
Variables	4	5	6	7	8
1 Number of centers participating in the trial					
2 Number of Location Countries					
3 Phase					
4 Enrollment target					
5 Previous publications	.117				
6 Tenure Status	.039	.076			
7 Past structural holes	.012	.023	-.184		
8 Cumulative Size of clinical centers participating in the project	.225**	.034	.065	-.038	
9 Years of operations of the firm sponsoring the project	.108	.009	-.119	.029	-.058

†Coefficients are significant at  $p < .10$

\*Coefficients are significant at  $p < .05$

\*\*Coefficients are significant at  $p < .01$

\*\*\*Coefficients are significant at  $p < .001$

**TABLE II**

<b>First Stage Regression (Endogenous Variable: Current Centrality)</b>		
<b>Variables</b>		
<b>Controls</b>		
<i>Number of Location Countries</i>	-0.00	(.00)
<i>Number of centers participating in the trial</i>	.00	(.00)
<i>Enrollment target</i>	.00	(.00)
<i>Years of operations of the firm sponsoring the project</i>	-.05	(.04)
<i>Cumulative Size of clinical centers participating in the project</i>	.06*	(.03)
<i>Phase</i>	-.01	(.02)
<i>Previous publications</i>	.00	(.00)
<i>Tenure Status</i>	.05*	(.02)
<i>Constant</i>	.56***	(.10)
<b>Instrumental Variables</b>		
<i>Past structural holes</i>	.10+	(.06)
<b>R-Squared</b>		<b>0.22</b>
<b>F-statistic</b>		<b>2.81*</b>
<b>N</b>		<b>60</b>

**TABLE III**

<b>Second Stage Regression (Dependent Variable: Number of citations)</b>		
<b>Variables</b>		
<b>Controls</b>		
<i>Number of Location Countries</i>	1.62*	(.74)
<i>Number of centers participating in the trial</i>	.20*	(.09)
<i>Enrollment target</i>	.02***	(.00)
<i>Previous publications</i>	.09	(.05)
<i>Phase</i>	1.50	(.00)
<i>Constant</i>	26.82	(22.74)
<b>Independent Variable</b>		
<i>Current Centrality</i>	-49.79+	(30.47)
<b>R-Squared</b>		<b>0.94</b>
<b>F-statistic</b>		<b>144.84***</b>
<b>N</b>		<b>60</b>

†Coefficients are significant at  $p < .10$

\*Coefficients are significant at  $p < .05$

\*\*Coefficients are significant at  $p < .01$

\*\*\*Coefficients are significant at  $p < .001$

Robust, heteroskedasticity-adjusted standard errors in parentheses

TABLE IV

Regression (dependent variable: Number of Citations)

Variables	Model 1		Model 2	
<i>Centrality</i>	-167.94†	-92.29	-417.24†	(-222.96)
<i>Regulation Constraints</i>			352.48†	(190.94)
<i>CentralityXRegulation Constraints</i>			-458.03†	(248.28)
<b>Controls</b>				
<i>Number of Location Countries</i>	2.82	(1.76)	2.79	(1.75)
<i>Number of centers participating in the trial</i>	.02	(.24)	.02	(.23)
<i>Enrollment target</i>	.01***	(.00)	.01***	(.00)
<i>Previous publications</i>	.37	(.25)	.36	(.24)
<i>Phase</i>	1.90	(2.77)	-.05	(1.77)
<i>Constant</i>	113.06†	(67.18)	305.53†	(168.15)
<b><i>R-Squared</i></b>	0.46		0.51	
<b><i>F-statistic</i></b>	86.42***		44.76***	
<b><i>N</i></b>	105		105	

† p &lt; 0.1

\* p &lt; 0.05

\*\* p &lt; 0.01

\*\*\* p &lt; 0.001

Robust, heteroskedasticity-adjusted standard errors in parentheses