Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Organizations as Social Cognitive Systems

Autor:

Juan Antonio Candiani

Directora:

Neus Palomeras Vilches

Departamento de la Economía de la Empresa

Getafe, mayo de 2015

Universidad Carlos III de Madrid

PH.D. THESIS

Organizations as Social Cognitive Systems

Author:
Juan Antonio Candiani

Advisor:
Neus Palomeras Vilches

BUSINESS DEPARTMENT

Getafe, May 2015

Acknowledgements

I would like to express my gratitude to my advisor Neus for her guidance, continuous support and patience above all. I am also thankful for the occasional but valuable advice from other professors at the Business Department, and for colleagues and classmates who have been not only amazing friends but also an endless source of knowledge and constructive commentaries. And beyond the academic field, I am infinitely thankful for my family and my friends that have always encouraged my desire of being a better person. Finally, I would also like to express my immense gratitude to the University Carlos III de Madrid for having given me the opportunity of researching inside its warm walls.

Abstract

This Thesis focuses on how organizations deal with complex knowledge by articulating the collaboration of their members. Three chapters address different aspects of this idea. The first examines the use of social networks for observing knowledge recombination and its implications for the innovation literature. The second chapter analyzes how knowledge of different levels of similarity can be combined throughout different types of social configurations. Specifically, multidisciplinary and redundant structures are operationalized and assessed in terms of different cognitive abilities needed for innovating. Finally, the last chapter uses innovations as a natural test for capturing the cognitive capacity of organizations for dealing with complex knowledge. This capacity is then related to how organizations decompose their knowledge bases into simpler structures.

Resumen

Esta Tesis se centra en cómo las organizaciones acceden a conocimiento de alta complejidad articulando la colaboración de sus integrantes. Sus tres capítulos analizan diferentes aspectos de esta idea. El primero examina el uso de redes sociales para observar la recombinación de conocimiento y sus implicaciones para la investigación de la innovación. El segundo capítulo analiza cómo conocimiento de diferentes niveles de similitud puede ser combinado por diferentes tipos de configuración sociales. Concretamente, estructuras multidisciplinarias y redundantes son operacionalizadas y evaluadas respecto a las diferentes habilidades cognitivas requeridas para innovar. Finalmente, el último capítulo usa las innovaciones como un test natural que captura la capacidad cognitiva de las organizaciones cuando se enfrentan a conocimiento de alta complejidad. Esta capacidad es luego relacionada con cómo las organizaciones descomponen sus bases de conocimiento en estructuras más simples.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This doctoral thesis explores the idea of organizations as social cognitive systems by focusing on their capacity for dealing with problems of large dimensions. Firms deal with all kind of problems on daily bases and it is their ability to solve them properly that make them succeed. Among them, innovating offers very interesting insights. Understood as knowledge generation based on previous one, innovations usually require the involvement of many people in their development as the scale of problem is too large for lonely inventors. Organizations generate the social space where complex knowledge is processed for that purpose. These two dimensions, the social and the epistemic dimension, describe the collective cognitive system as if they were the neural network and the cognitive map of a human brain. Throughout its three chapters, this thesis analyzes organizations as the interaction of social structures and knowledge by focusing on their innovation activity.

Chapter 2 triggers the question of whether the social space and the knowledge space can be considered as isomorphic. The increasing use of social network analysis applied to the innovation field has largely relied on this crucial but implicit assumption. Social distances are generally taken as an indicator of knowledge heterogeneity among network's component. This assumption allows translating properties from one space to another; something can be very useful when studying innovation as combination of knowledge. If knowledge is decomposed, held by many people and collectively combined, the characteristics of the social structure can describe characteristics of the knowledge structure. However, this assumption has apparently been overlooked by the main literature. This needs not only to be revised in order to valid empirical results or theories based on this premise, but it also opens interesting research veins. Because of these reasons, Chapter 2 begins the analysis of organizations as cognitive systems by exploring the social and the knowledge structure of a University.

Across the three chapters of this doctoral thesis, innovations are considered as novel combinations of prior knowledge. Instead of focusing on other dimensions, this definition stresses the genesis of an innovation. It postulates two major ideas. First, any innovation is generated on previous ones. Second, innovations are new combinations. This perspective allows including other phenomena beyond technology. It applies to knowledge generation in general. Academic research can be analyzed as innovations under this perspective. Instead of inventing, academic researchers combines previous knowledge in novel ways in papers and books. Chapter 2, then, opens the analysis of innovations as collective processes analyzing the academic research of an University. Notwithstanding the modest scale of the empirical analysis, this chapter untangle complex information for independently capturing both the social and the knowledge space that describe three research departments of this university.

Considering a three-year period, Chapter 2 uses three different data sources of analyzing the structures of the two spaces. For describing the knowledge space, all papers, books and working papers produced by this University are analyzed both in terms of authors and co-authorship, and in terms of backward citation for content similarity. With regards to the social structure, Chapter 2 uses the network of interaction of institutional e-mails and a direct survey collecting information about different kind of social relationships among researchers: those based on technical advice, those based on information regarding the university but not the research activity, and those of a personal nature. The small scale of this analysis is the characteristic of this chapter that allows collecting detail information on the social structure. With this data set, different metrics are developed in order to capture distances among researchers in both spaces. These distances generate a metric space for the social and the knowledge space, distances that are statistically compared using different techniques. Empirical results show that the social distances are not a simple monotonic transformation of knowledge distances within the three research departments considered with only one exception: relationships based on technical and professional advice. Although the statistical analysis shows a significant positive relation, estimated coefficients are not strong enough to support a monotonic transformation among distances in both spaces. These results show that social structure cannot be simply taken as indicator of knowledge or information heterogeneity. Broadly extended ideas as centrality on social network, boundary spanners, weak ties or structural holes among others, lose strength when considering these results. Even though social structures are a powerful tool for analyzing innovation since they can describe the collective processes of knowledge combination, they should be complemented by considering the other natural part of the puzzle: the knowledge embedded in their members. By questioning the assumption of isomorphism between the social and knowledge space, an interesting question emerges: what happens if social and knowledge spaces are not isomorphic? Not only some previous research should be revised but it opens a fruitful field.

Chapter 3 explores the interesting consequences of dropping this assumption. If social proximity of people does not describe the similarity of knowledge they poses, two scenarios may be observed. On the one hand, a group of people may be characterized by being socially close but the knowledge embedded in those persons very heterogeneous. In other words, this would be a group characterized by short social distances but long knowledge distances. This case can be illustrated by a multidisciplinary team whose members closely work by assembling diverse expertise. On the other hand, different persons may be experts on the same knowledge areas but do not know each other and have no people in common. They would be closely located in the knowledge space but distant in the social space. If these people work for the same organization, they could be seen as duplicated experts since they operate independently although their overlapping of knowledge. These two opposite cases describe well differentiated socio-knowledge configurations that impacts on the innovation activity. This relation comes natural when considering innovation is considered as recombination of previous knowledge, and therefore, influenced by how the social structure combines or disaggregates it. Furthermore, when applied to the innovation field, this benchmark allows studying a very interesting concept. If the task, role or function of an inventor is to research on those areas she or he is expert on, knowledge expertise can be identified with function within an organization. Then, when socially distant inventors working for the same organization are expert on the same areas, and therefore, they research on the same topics, they can be said redundant. Thus, redundancy could be stated as opposite to multidisciplinarism or cross-functionality.

As different as they are, these two structures are expected to influence on innovation. However, their intrinsically different nature should manifest in different kinds of innovation. Chapter 3 introduces a novel distinction of innovations for that purpose. These are classified according to the cognitive abilities which are necessary to generate successful new configurations of previous technologies. Considered assembles of different pieces of knowledge or technology, inventions generate two different cognitive challenges according to the level of interdependence of their components. When components are highly interdependent, changing one of them implies affecting the performance of many others like a moving a single piece in a Rubik Cube. On the other extreme, when components have almost not interdependence, changing one piece has almost no effect on others pieces contribution to the overall performance of the invention like if they were Lego pieces. Highly interdependent technologies need different cognitive abilities to find new configurations than when dealing with low interdependent technologies. While the former needs systematic processing of the many variations, the latter needs disruptive thinking. Thus, innovations are analyzed according to these two different cognitive challenges.

Chapter 3 empirically assess the relation of the two opposite socio-knowledge configurations, cross-functional and redundant structures on the two types of innovations according to the cognitive processes needed for generating them. The analysis of patents generated in the semi-conductor industry examine whether cross-functional structure foster disruptive thinking while redundant structure facilitate systematic processing. Results support these relations and challenge a very rooted idea on management: the idea of redundancies as duplicated costs that must be eliminated in the name of efficiency. Redundancies within organization may offer an effective way of dealing with extreme interdependent problems. This chapter's contribution to the innovation field relies on proposed theoretical framework and operationalization of these concepts.

The social structure combined with the distribution of knowledge across the members of an organization proves to be crucial when innovating. Since the complexity of knowledge can easily outruns the individual capacity, it is natural to think knowledge as being combined throughout social collaboration. Since their beginnings, human communities have used the strategy "divide and conquer" for dealing with complex problems. Societies formed on the bases of specialization and collaboration, which allowed the continuous expansion of the knowledge they can embrace. And within societies, organizations push this strategy further. By having common goals and a shared identity, organizations foster coordination, communication and learning of their members. Organizations, then, can be said to have or be a collective intelligence capable of solving much more complex problems than any single member could. This idea of firms as information-processing machineries, although very intuitive, has been never empirically captured.

Chapter 4 proposes a method for observing this intrinsic capacity of organizations for solving highly complex problems. Human intelligence, as measured by standardized tests, is basically a metric that captures success of solutions controlling for the intellectual difficulty of the problems. Innovation offers the same two independent dimensions for observing this capacity on firms. On the one hand, when patented, the success of innovations can be measured by forward citations. On the other hand, the intrinsic difficulty of being able to succeed with certain invention can be measured by the level of interdependence of its components. Previous research has shown that interdependence of its components makes innovations difficult to perfect but allows some rare but highly successful configurations. Therefore, when controlling for the intrinsic difficulty that interdependence generates, the success of innovation would reveal which firms are better suited for dealing with such complexity.

The organizational capacity to find rare but very successful combinations in these contexts is then related with the way organizations simplify complexity. It is well known that human brains are expert in finding patterns in any phenomena. The relation between understanding and capacity for simplifying a phenomenon is almost a definition: the way we understand *is* by simplifying. Our mind structures concepts and ideas in hierarchical near-decomposable structures. That way, those elements that are detected as similar or strongly connected are grouped within groups, while relations among groups are almost negligible. This intrinsic and defining characteristic of our way of understanding can be translated to organizations. As they research and innovate over the years, organizations explore and combine different technological classes. Some classes are used more than others; some combinations of classes are more frequent than others. Organizations, then, reveal what types of technologies they explore and how they understand relations among them. In other words, when innovating, firms reveal how they understand technology and how much they are capable of simplifying it into meaningful structures.

By innovating, organizations reveal their capacity for dealing with complex phenomena *and* how they decompose the technology field. Chapter 4 empirically assess the relation between these two. It is natural to expect a positive relation between the ability for simplifying complex problems and the capacity for solving them. Using two different approaches, Chapter 4 find empirical results that support the expected relation. The negative relation between complexity of technologies and expected success of inventions does not hold when the "intelligence" of the organization is considered. Consequently, this chapter stands behind a broadly accepted idea: the expected success of an invention on very complex technologies is expected to be high *if* you have the intellectual capacity.

Furthermore, the methodology used in this chapter explores a way of capturing an idea that many managers would want to measure for assessing their decisions. Hiring certain profile of workers, getting rid of subsidiary firm, merging two business units, or changing the head of some department could have substantial consequences on the "collective intelligence" of the firm. Thus, this chapter can be seen as a step towards the operationalization of this concept. By stating a parallelism with persons, the capacity of organizations to deal with complex problems could anticipate their performance in the long run.

This Thesis explores innovations as complex knowledge that is combined and processed throughout social structures. Its three chapters show that innovations can be serious cognitive challenges that require strong social collaboration, a not trivial goal that may be only achievable by organizations as social meta-identities of large groups of people. The capacity for understanding complex problems and the configuration of how knowledge and people are combined depict organizations as social cognitive systems.

Chapter 2

The Clash of Social and Knowledge spaces

The Assumed Isomorphism Under the Hood

Abstract

This chapter explores the very concept of innovation as novel combinations of knowledge. Innovators or researchers are located in a knowledge space according to *what* they know and in a social space according to *who* they know. The former space is built upon the concept of compatibility as a measure of distance between different pieces of knowledge, while the latter is represented by a social network. The interaction of these two spaces determines the potentiality of innovating. However, so far, most of the innovation literature has assumed them as isomorphic. Specifically, the use of social networks in this field takes social distances as a measure of knowledge diversity. Using data on publications and working papers, e-mail activity and a sociometric survey that captures three kind of relationships (technical, institutional, and personal), three related research departments of a University are empirically analyzed for assessing this assumption. Both their knowledge space and social space are independently reconstructed using different measures for then testing the correspondence between distances. Results show that the only weak but statistically significant relation is between knowledge similarity and social distances by considering those relationships based on professional consultation (technical). The other measures used to build the social space, although broadly used across the literature, fail in capturing knowledge similarity. These results suggests that the social space cannot be used straightforward to analyzed the knowledge space as most of the innovation literature implicitly does. The knowledge embedded in social networks must be considered before analyzing their structure and their influence on innovation. Disentangling these two spaces may offer rich insights on the problem as Chapter 4 explores.

2.1 Introduction: Innovation as a social outcome

During the last decades there has been an explosion of social network analysis in many scientific areas and the research on innovation has not been the exception (Phelps et al., 2012). As in any other activity we practice as social beings, innovation is certainly constrained by the social surrounding (Granovetter, 1985) in the sense that it cannot be reducible to the study of individuals but it must be understood by considering the collective. Rich friends might support your ideas, highly educated colleagues of your university can help you developing a project, or your firm's neighborhood might stimulate your innovation performance. However, differently to other areas, in the innovation field the social network analysis might offer more than just considering its social dimension. As I will explain in the following paragraphs, given some contextual conditions, social networks' structures of innovators or researchers may mimic knowledge structures. Such property makes social networks the perfect candidate to study how innovation is built upon knowledge.

Usually attributed to Schumpeter (1939), one of the most famous definitions understands innovation as novel combinations of knowledge. Instead of stressing on other aspects, this definition focuses on their theoretical genesis which is very difficult to unfold into observable phenomena. No one can keep track of those mental processes that combine knowledge and ideas inside an innovator's head. However, the innovation literature might have found in the social network analysis an attractive solution for the operationalization of the concept. In order to observe how knowledge is combined, the scale of the problem must be changed. Instead of thinking on a lonely inventor reading books and using a board for designing new ideas, we should think her as incapable of dealing with the necessary knowledge to innovate, and thus, forced to collaborate with others (Jones, 2009). This assumption is very realistic given the widening gap between the average person's intelligence and the intrinsic complexity of the knowledge we are immersed in nowadays. Not only in terms of what we should know to innovate but also as the combinatorial problem that the searching process represent (Sorenson et al., 2006; Fleming and Sorenson, 2001, 2004). Looking for the best possible combinations easily exceeds the capacity of a single person. As a consequence, the dependence on intellectual collaboration is not an option but an unavoidable necessity that can be taken as the starting point for any analysis on innovation.

When the complexity of knowledge outruns the capacity of individuals, they must specialize in small knowledge subsets and collaborate with other specialists in order to embrace wider areas (Kogut and Zander, 1992, 1996). Collaboration is the way in which knowledge interacts through people. Sociology has a long record of considering social ties as channels or conduits where knowledge flow through (Granovetter, 1985; Burt, 1992; Uzzi, 1996; Podolny, 2001; Owen-Smith and Powell, 2004a), what constitutes the cornerstone of any social analysis of knowledge creation. By definition, social interaction is built on communication, and therefore, it is capable of transmitting information from one party to another. Thus, controlling for the pertinent characteristics of people, relationships and content of the transmission, a social structure of interactions should be capable of describing *how* knowledge is routed among them.

Since people are forced to interact in order to deal with the knowledge they cannot manage by themselves, the social network analysis offers a powerful tool to capture the resulting combinations of knowledge. When people need to specialize in infinitesimal areas of knowledge innovating demands many participants. Then, by observing the social network they form we could have a clear picture of how knowledge is being combined. That is why, when applied to this field, social networks offer more than a social analysis. They may *mimic* the knowledge structure and therefore they could describe how knowledge is generated. This outstanding property is what makes the social network analysis the rising star in this field (Phelps et al., 2012).

However, all that glitters is not gold. The innovation literature may have pushed this relation too far. When social networks are used to understand innovation as combinations of knowledge, a direct correspondence between social proximity and knowledge similarity is generally assumed. If we think innovators as located in a social space according to *who* they know and in a knowledge space according to *what* they know, what the literature generally assumes is an *isomorphism* between these two spaces. Even though both spaces are not independent since people might build relationships with similar people or the other way around, the assumption of an isomorphism might be too simplistic.

Assuming that the social and the knowledge space are isomorphic, so far, most of the claims made by the literature linking innovation and social networks are corollaries of this idea. As an example, from the seminal work of Granovetter (1973) many scholars have analyzed social networks at the relational level, i.e. focusing on characteristics of the social tie and how it influences on its ability to transfer knowledge (Uzzi, 1996; Nahapiet and Ghoshal, 1998). It is generally accepted that while strong ties are able to transmit complex and sensitive information and weak ones cannot, the latter have the ability to bring fresh and distinct information from further groups (Bouty, 2000; Hansen, 1999; Reagans and McEvily, 2003; Perry-Smith, 2006). As it can be observed, this is a logical consequence of the isomorphism: weak ties, in a social network analysis are considered as direct but distant edges. Thus, as distant social relations, they are able to bring distant (different) knowledge. In both cases, the correspondence between social and knowledge proximity explains the ability to innovate.

At a structural level, social networks applied to innovation are also treated as isomorphic to the knowledge space. When considering particular positions in a network, the cohesion criterion indicates how rich and diverse the knowledge a person can have access to is (Phelps et al., 2012). For example, it is stated that central players get the greatest amount of information in the network since they are located at the closest position to the rest of the members and this affects their ability to innovate (Granovetter, 1992; Newman, 2001, 2010; Tsai, 2001; Tsai and Ghoshal, 1998). On the other hand, boundary spanners are also influenced by their position in the network: they are far from the center of their group but closer to the exterior, what it means that they are more permeable to foreign knowledge and less attached to internal one (Tortoriello, 2005; Tortoriello and Krackhardt, 2010).

One of the cornerstones of the social network analysis is also based on this isomorphism.

Burt (1992) proposed the concept of the structural hole as the situation when a social actor has non-redundant contacts, and therefore, it is expected to have access to non-redundant information. Again, a person who is located at a structural hole will be closer to those sides that she is bridging, and therefore, closer to the knowledge they have. Furthermore, when the structural hole is broken by closing triads, the variance of distances among nodes decreases as the differences in knowledge between people do (Tortoriello and Krackhardt, 2010).

Those researchers that use a more holistic approach considering the whole network structure are not the exception concerning this issue. In structural analysis, the probability of knowledge transfer between individuals is considered to diminish as the path length between them increases (Phelps et al., 2012; Reagans and Zuckerman, 2001; Reagans and McEvily, 2003; Singh, 2005). Therefore, for instance, density of social networks is considered to foster common knowledge and shared meaning as it decreases the diameter of the network and the heterogeneity of distances among nodes (Kogut and Zander, 1996; Tortoriello and Krackhardt, 2010; Woodman et al., 1993). Small-world structures, as another example, are also expected to achieve an optimal equilibrium for innovation because of their clusterization combined with a small diameter due to the presence of bridges that jump from cluster to cluster (Fleming and Waguespack, 2007). As it forms a near-decomposable structure, i.e. it achieves a balance between concentration and breadth, small-world structures are said to be optimal for innovation since near-decomposable structures of knowledge are optimal (Levitan et al., 1999; Yayavaram and Ahuja, 2008).

Without trying to be exhaustive, the former examples show that the mainstream research on knowledge creation and social networks directly assumes an isomorphism between social and knowledge structures. Considered a consequence of the increasing gap between knowledge complexity and average human intelligence, collaborative relations are taken as manifested knowledge compatibility. Researchers has to reduce their area of expertise so much that they can be considered as fundamental pieces of the entire body of knowledge embedded in society. Although so far social networks have been used in a broad perspective to explain mainly knowledge heterogeneity, the assumption of isomorphism must be revised. It might be the source of conflicting results in the empirical analysis as Phelps et al. (2012) shown in their exhaustive literature review on the field, or a simple neglect about the distinction between social and human capital (Nahapiet and Ghoshal, 1998).

This chapter does not pretend to discuss how social structures affect knowledge transmission or creation but it seeks to challenge the isomorphism taken for granted by many papers that research on innovation as socially driven. Are social spaces isomorphic to knowledge spaces? So far, the literature has overlooked this question. Should the previous question be negatively answered, a large part of that research should be revised.

In the following sections I discuss the theoretical construction of both spaces innovators are embedded in when creating knowledge. Afterwards, I perform an empirical analysis on the research activity of three departments of an University devoted to statistics, economics and business. In order to do the latter, the knowledge and the social space must be independently identified. On the one hand, for describing the knowledge space, I collect information about papers and working papers of all researchers working for those departments. Analyzing backward citations I measure similarity of knowledge content among papers and researchers during that period. On the other hand, three alternative approaches are used for capturing social distances. First, I used the institutional email network among researchers. Second, a direct survey asked about three kind of social relations: those based on technical advice and professional consultation, those based on conversations about university-related matters different from research, and those based on personal and more intimate nature. Third, co-authorship was explored as a way of capturing collaborative ties. Results shows that only social distances of those relationships based on technical consultation have a significant but weak correspondence with knowledge similarity among researchers. The co-authorship criterion, even though analyzed, is discarded due to the lack of variance on the data set. The empirical work in this chapter performed suggests that social networks should be carefully used when analyzing information flow and knowledge combination in the study of innovations. Given that social structures describe the flow of knowledge, the knowledge embedded in its components must be considered as well.

2.2 The two spaces

The study of innovations and social networks has been based on a major assumption. As explained in the introduction, when knowledge creation is understood as recombination of prior knowledge, the social dimension might help describing much of this process. Since social relationships are based on information exchange, the structures they form could describe how knowledge might be combined along its transmission across people. This reasoning has led scholars to implicitly treat social closeness as knowledge similarity, and thus, social structures as isomorphic to knowledge structures, the ground where much empirical research on innovation has been conducted on.

As any assumption, if fake, the subsequent logical reasoning does not guarantee the truth of its deductions. Assumptions must be tested and corroborated in order to state their coherence with the empirical experience. The literature on knowledge networks has never thrown doubt on this cornerstone assumption and this is what this paper intends to do.

A proper examination of the assumed isomorphism between the social and knowledge space should be conducted in order to state the validity of the social network approach for studying innovation as combination of knowledge. This is what I will endeavor to in this section. As I am going to challenge this assumption where much literature has built and tested different hypotheses, a clear distinction must be made between the social and the knowledge space. To judge whether they are isomorphic implies considering both spaces as ontologically independent phenomena, and therefore, they should be independently constructed in order to study their similarity.

However, the separation of both phenomena might not be trivial. While social closeness can be captured by relatively simple means, the similarity among what persons know represents a bigger challenge because of the inherent difficulty to define knowledge. Social relationships are built on repeated interaction and therefore, they are potentially observable. On the other hand, it is almost impossible to define knowledge without triggering an infinite debate about its nature. As a consequence, there are not simple means of observing what people know and how similar regarding to what others know.

In order to unfold what has been assumed as identical, independent measurements and a proper data set are required. As long as innovation is considered as knowledge recombination through people interaction, the data set must contain a group of people collectively generating knowledge for capturing their collaborative activity and their knowledge background. For doing so, this chapter analyzes the activity of three research departments of a Spanish university. Considering academic research similar to innovation as both recombine knowledge in order to generate new one, the two spaces are built as analytic tools to explore whether there is a resemblance strong enough to be taken as isomorphic. Even though the scale of the empirical exercise is modest, the technique is suitable for larger cases.

2.2.1 The knowledge space

Describing innovation as happening in a knowledge space does not have any ontological commitment but it is only a conceptualization based on a human cognitive dimension. It tries to capture an intuitive way of thinking not only knowledge but any phenomena we perceive. This is what Simon (1962) referred as a hierarchical way of understating. Our cognitive system is expert on finding patterns, on discovering similarities and dissimilarities between objects that end up in hierarchical classifications. The abstract perception of knowledge is not the exception. In the academic argot, expressions such as 'fields of science', 'research gaps', 'branches of knowledge' or 'research area' among others illustrate this. This section will discuss some explicit methods of capturing this perception that can be translated into a spatial concept.

In order to analyze whether there is an isomorphism, a knowledge space must be constructed independently from the social space. This space would describe distances among knowledge content of different knowledge supports. In the case of this paper, researchers would be located in the knowledge space according to what they know such that their distances reflect the similarity of their expertise.

However, it is not trivial to capture something as a knowledge distance since the concept of knowledge itself is very hard to define and operationalize. The definition of knowledge, though, might be *not* necessary to capture differences among knowledge content either in people or other supports. When studying knowledge creation understood as knowledge recombination, the goal of defining knowledge proximity is functional to analyze how likely or potentially combinable are what people know. Across the innovation literature there is this idea of an optimal heterogeneity of knowledge to obtain a succesfull innovation. Redundant knowledge brings nothing new while totally different and unrelated knowledge impedes any attempt of combination. Therefore, something in the middle may achieve the perfect breeding ground for innovations, gathering knowledge not so much heteregenous neither much homogeneous (Yayavaram and Ahuja, 2008). The idea of potential compatibility is the one I want to capture to understand innovation and it can be thought in terms of distances.

That is why I consider knowledge distances as manifested compatibility. Those knowledge elements that were combined into new knowledge are compatible and the other way around. The compatibility does not refer to inherent condition to be combined but to a manifestation of such. For instance, if an architect consults expert civil engineers and geologists in order to build the foundations of a skyscraper, their areas of expertise manifest to be compatible to develop new knowledge. Knowledge needs not to be defined to manifest their structure.

Based on this approach, several methods can be used to study similarity between knowledge content in different supports. Probably the most straightforward approach to capture differences in knowledge is its explicit classification. This might be the oldest method, performed from ancient philosophers to modern scientific journals. Taxonomies as JEL (Journal of Economic Literature) or AMS (American Mathematical Society) codes for papers in economics, mathematics and statistics, or the USPC (U.S. Patent Classification) for US patents are examples of this approach. These classifications seek to divide an entire set of possibilities according to some criteria such that those objects considered as significantly different are classify with different labels.

Based on the previous technique of pre-established taxonomies, Yayavaram and Ahuja (2008) proposed an interesting way of analyzing knowledge structures. They considered a patent as a link between those technological classes the patent was classified into. Then, taking a firm's portfolio of patents it could be seen how often two technological classes were linked, and thus, how the firm *structures* knowledge. This approach could be generalized to the entire universe of patents to have a complete network that pictures how technological classes are related. Heavy linked categories would indicate highly compatible technological categories and the network of categories would depict distances between unrelated categories. The same conceptualization would apply for academic papers with an equivalent classification like the JEL code.

Even though this approach is very simple to understand, it lacks of enough flexibility in evolving contexts. Categories are preset by experts seeking to divide the entire universe of possibilities in representatives groups. But when science or technology evolves towards higher complexity, categories eventually become obsolete, insufficient, or not enough fine-grained to describe the internal subdivisions of previous categories. Imagine that the category 'automobile' was pretty precise when the car was invented, but few decades after, the spectrum of automobile vehicles had increased so much that it needed many new subcategories to capture the variety. Equivalently, JEL or USPC categories will eventually fall short in describing the space where papers or inventions are located, and the entire taxonomy must be restated with the associated cost of changing. That is why other approaches have been proposed.

When analyzing academic papers or patents, there is an alternative way describing a knowledge space that dispenses with categories. Either papers or patents explicitly refer to previous papers or patents they were based on. Then, we can let knowledge creation to happen and keep track of these combinations. What is combinable will be eventually combined and the network of combinations, throughout time, will draw clusters of suitable or successful technologies or knowledge (Fleming and Sorenson, 2001, 2004; Sorenson et al., 2006). These clusters should converge to the categories used in the previous approach although not necessarily. This way of describing the technological or knowledge space evolve ignoring failed research, creating new space for successful branches, and allowing any fine-grained description as it is needed.

Networks of citations are the ones that keep track of combinations in the generation of knowledge. This analysis relies on the explicit references a paper or a patent does regarding previous work it was built upon. The network of citations describes distances between papers (or patents) as if they were located in a knowledge space and, as such, it constantly reshapes as science or technology evolve.

A citation indicates that the citing paper uses the knowledge of the cited paper as its component (Fleming and Sorenson, 2001; Martinelli, 2011). The set of a paper's referenced papers manifest to be combinable and should be closely located in a knowledge space. Since researchers aim to contribute with new knowledge, and therefore, to show their contribution, they must accurately identify the knowledge frontier by the references they used. Thus, they should be a precise criterion to indicate the components of research or innovations, even more when considering that references also double checked by independent experts when papers are published or patents are evaluated to be patented.

Summing up, by using a citation approach, I can propose a measure of knowledge content similarity based on manifested compatibility in order to build a a knowledge space using as a unit of knowledge a scientific paper or patent. For the sake of the analysis, once a measure of compatibility is defined as a distance, I can propose a metric space that generates the whole set of distances between any considered knowledge component. As it describes knowledge, I refer to this space as a knowledge space. There I could locate people according to their area of expertise, such that people whose expertise often complement in research or innovations are epistemically close, and people whose specialties are hardly combinable would be far located. This epistemic space would be able to explicitly capture the intuitive idea we have about the structure of knowledge.

2.2.2 The social space

In the innovation literature social networks are used for explaining process of knowledge recombination. Since knowledge flows through communication between people, different structures and locations within the network affect the innovative performance. As evident and intuitive this idea might seem, social networks have been forced to describe more than they can. Explaining knowledge recombination when it is ignored how knowledge is spread across people turns out to be impossible. That is why, for bypassing this problem, social networks are assumed to describe knowledge similarity among its members by social closeness. Since in this paper I intend to assess the assumed isomorphism between both structures, I must consider the social dimension.

Knowledge creation is intrinsically a social process, particularly when facing high complex knowledge, since people specialize and rely on others' knowledge as the only way to learn, process and create new knowledge. Social relationships work as wires that connect people's intellects in a very complex landscape of interactions. This web of knowledge conduits is what conforms the social space where innovators are embedded in.

The idea of a social space is pretty intuitive. Everyone has a perception of social closeness to others, not only from those we know directly or indirectly but, as Milgram (1967) showed in his famous research, from unknown people. The social space as the network of relationships among persons would describe social distances understood as the effort of routing knowledge from one person to another, or the likelihood that knowledge reach one person from another. In the social space, friends, acquaintances and colleagues would be closely located, whereas distantly related people would be further. As the knowledge space, this space is also built on a relational approach, but instead of compatibility of components the social space uses social ties to describe direct distances between people, and indirect ties to measure the rest of the distances.

Since social proximity increases the chances of sharing information, the innovation literature has devoted to capture social networks among innovators in order to study how their access to knowledge affect their performance. However, capturing a social network is not trivial not only because it is a dynamic structure that constantly changes but also because social relationships are difficult to define and capture. Different ways of registering a social network have been used along the literature.

Some researchers have used the co-authorship criterion to detect social structures. Co-authorship considers that two researchers (or inventors) are linked if they published (or registered) together a paper (or patent). It has been broadly studied across the literature on innovation since it describes social collaborative ties and therefore flows of knowledge between researchers (or inventors) (Newman, 2001, 2004; Pepe, 2011). The underlying logic here is that if two researchers work together, they necessarily have a collaborative link built upon communication.

However, there is another reading of what a link a co-authorship network is: knowledge proximity. Two researchers that work on the same paper necessarily share knowledge. And, since time and resources are scarce, and a professional researcher must publish in the area she thinks will have the best chance to publish, then, scientists will work on papers related with their best expertise. Hence, co-authorship means that the authors' areas of expertise overlap and complement. Summing up, the network of co-authorship can be seen as a knowledge space, where proximity captures similarity in what researchers know. This ambiguity between the social and the knowledge spaces the co-authorship network captures lies at the core of the commonly assumed isomorphism between both spaces. That is why I will ignore this approach to focus on other alternatives.

The co-authorship criterion may be chosen for researchers because of its methodological simplicity but because of its ambiguity, other techniques must be explored. Among others, two general methods are broadly accepted to capture social networks: sociometric surveys and electronic communications.

On sociometric surveys, a set of questions intend to make people reveal who and how they are related with others. Among the different problems a survey might face, when devoted social networks it deals with a particular issue: the rate of response. The absence of a tiny percentage of individuals' answers can radically affect the entire topology of the social network.

The second method is based on communicational electronic data sets. During the last years, the use of these data sets have arisen as an alternative source of information for capturing social networks. The increasing availability of electronic records of communicational activity along with major computational power triggered a new wave of large scale social network analysis. Among other advantages, these data sets are easier to collect, are exhaustive since they do not depend of the collaborative attitude of the surveyed person, and they are far more detailed than a survey can be specially with longitudinal records. However, they are not exempt from doubts. Some researchers question how new means of electronic communication could affect the way we relate with people (Treviño et al., 2000), others question how accurate they are reflecting surveyed-based social networks (Grippa et al., 2006; Lex et al., 2011). Beyond these debates, networks of electronic communication are broadly used to describe social networks (Gloor et al., 2003; Guimera et al., 2003, 2006a; Kossinets and Watts, 2006; Wellman et al., 1996).

In this paper I will use these two approaches in order to reveal social closeness among researchers towards a comparison with knowledge similarity of their area of expertise. As I will explain later on this research, the electronic data set of communicational activity will be used to calculate social distances among members of the sampled University while the survey-based approach will be considered to locally check knowledge distances. The general idea is to evaluate researchers in a social space as the one that determines how likely they collaborate and share knowledge.

2.3 Empirical approach

The goal of this paper is to empirically assess the assumed correspondence between social and knowledge distances that has lead the literature on studying knowledge creation by using social networks. For that purpose I intend to independently construct both the knowledge and social structure of academic researchers working for a university in order to analyze the possible presence of an isomorphism.

The data set for this empirical analysis comes from the research activity of a Spanish University. Three research departments were considered: Economics, Statistics and Business Department. The three of them have successful indexes of scientific publication and research topics are expected to overlap. At the same time, there is much social interaction. Their buildings are closely located, not further than 100 meters in the same Campus, so direct interaction is quite common.

As research departments, scholars perform knowledge-intensive tasks. There is almost no other task rather than teaching and researching, but it is the latter the one that constitutes the main goal. Scholars aim to generate knowledge and because of that collaborative links are expected to happen. Indeed, when there is not any need of specific equipment such as laboratories or machinery, the essence of having scholars working under the same roof is fostering collaboration. In other words, these research departments constitute the perfect ground to study how knowledge is collectively generated since their main goal is such and social activity is expected to happen as an essential part of the process.

Furthermore, the research activity of these people manifests in academic papers that can eventually be published in specialized journals. As mentioned before, academic papers sometimes are classified by scientific codes to determine areas of study and they explicitly reference the previous work they are based on. Either way, they can provide much information about the knowledge activity within the research departments.

Based on the conception of an academic researcher as collaborating with others and creating knowledge from the recombination of what she knows and what colleagues know, I will assess whether the social network of researchers is isomorphic to the knowledge structure that characterize them, so the social network could be used to explain knowledge interaction ignoring the other one.

2.3.1 The knowledge space

In the previous sections it was discussed how to assess similarities (or distances) among researchers according to what they know. For doing this, I analyze academic papers the three research departments produced. I consider these papers as new knowledge generated by combining the one expressed in their bibliographic references. Furthermore, I also assume these papers to a manifestation of what their authors know and specialize into.

In order to describe the knowledge structure of the Economics, Statistics and Business research departments I collected information from their respective official annual records. These annual reports of research activities are lists of all published and working papers produced during a natural year. They are described by title, authors, and journal (in case of a published paper). This information was expanded by including JEL or AMS code (when possible), year, department and research area. Considering years 2008, 2009 and 2010, these records contain 606 papers.

Given that not all the papers but a minority of them were classified in JEL or AMS code, the use of this pre-established taxonomies of knowledge was discarded. The analysis of their content, then, was based on papers' references as indicators of previous knowledge the research was based on. The network of citations among papers can be considered as a self-organized system that describes the knowledge space. However, few data bases attempt to record the complexity of these patterns and none of them could be used for this paper. Although wide and well-known data bases of paper citations as the Web of Knowledge or Google Scholar can be consulted to analyze citation patterns, since some papers of this University are working papers and others were published in non-indexed journals, I had to call on other sources. Furthermore, even if a paper is registered in one of these citation bases as Google Scholar, most of the times they do not exhaustively account for all the references a paper has.

For these reasons I constructed my own data base accounting for *all* references declared by *every* paper registered in the official records during the three analyzed years. An algorithm was designed to store references from academic papers in PDF format identifying author(s), title and year of publication as three different alphanumeric fields. For the sake of simplicity, even though the data base involves three consecutive years, papers were considered as simultaneously produced, ignoring cross references inside the data base. As a result, the whole set were constituted by 606 papers with 19,028 references in total.

A paper's references indicate where its authors are exploring the knowledge space. So, in order to account for shared components that could indicate similarity, I sought for common references among papers. Since they were not indexed and they were stored as alphanumeric information, in order to find coincidences between every pair of paper's references I used text mining techniques. Combining the Levenshtein string metric for analyzing authors' name and title similarity, and checking with year of publication, coincidence of references were detected. As a result, the group of registered 606 papers collectively cite 12,365 papers as references. The citing group and the cited group work as a bipartite network since I ignore references within both groups. Differently to other citation analyses, this data set considers all papers and working papers produced by the University and all their references.

As a consequence of the way it was constructed, the data set only considers backward citation. Even though taking into account forward citation would improve the analysis, I ignored them since papers in my data base have been recently produced and not all of them have been published -what almost annihilates the possibility of being cited. Furthermore, when understanding knowledge creation as recombination of knowledge, backward citation clearly fits the idea of an author stating what she has combined in order to generate her new piece of knowledge. In other words, while backward citation can describe knowledge content of a paper while, forward citation describes how a paper is interpreted or used by other authors.

Then, in order to describe similarities about the knowledge content of papers, I relied on the concept of structural equivalence considering only backward citations. It would be expected for two papers with similar references to be researching on the same knowledge area. The more references they share, the closer their knowledge location should be. In the citation network, when two papers have identical set of references, they are *structural equivalent*. As logical consequence, they must be equally distant to all the rest of the nodes in term of geodesics and therefore, they would be located in the same coordinates in the non-euclidean space the network is describing. In other words, structural equivalence can be considered as the zero distance in the space associated to the citation network that I consider the knowledge space.

The property of being structural equivalent, though, is dichotomous: either two papers are structural equivalent or they are not. Because it is very unlikely to find two papers with identical references, a graded version of this concept has to be used to capture the entire range of papers which share references although not all of them. Generally, in network theory, similar nodes are expected to connect to the same set of nodes. This idea applies to citation networks as *bibliometric coupling* and it is generally assumed to depict similarity (Newman, 2001). It states that if two papers have common references, they research on similar areas.

Therefore, accounting for shared references between each pair of papers from the the database I used two measures of similarity. First, an standard measure of similarity, the *cosine similarity* which in large data sets converges to the *Pearson coefficient*. And second, one developed by myself that I will refer in this paper as *intersection of references*. If x and y are papers, both distances satisfy:

$$\begin{aligned}
&d(x,y) \geq 0 \\
&if\ d(x,y) = 0 \iff x = y \\
&d(x,y) = d(y,x) \\
&d(x,z) \leq d(x,z) + d(z,y)
\end{aligned} \tag{2.1}$$

As they observe these properties they can be used to build a metric space which I will consider as the knowledge space. After eliminating repeated papers between research departments and years, and discarding those papers that were isolated from the main core of papers in terms of shared references, the total number of papers contained in the data base of 3 years and 3 research departments is 497.

Cosine similarity

In network theory, when looking for a measure of node similarity only based on their pattern of connections, two measures are standard: the cosine similarity and the Pearson coefficient. Both of them analyze a node i's connections described by a vector of dimension n (being n the number of nodes in the network) where position j is 1 if the node is connected to node j and 0 otherwise.

The first proposes measuring the similarity between two nodes as the cosine of the angle that the two vector describe. The second measure, the Pearson coefficient, describes similarity between nodes as the linear correlation coefficient between the two vectors of connections. When the network is large enough, both measures converge which is why, only the cosine similarity will be considered for this research.

Under this approach, two structurally equivalent nodes would have *cos=1*. On the contrary, two nodes that do not share any common tie would account for cos=0. The cosine similarity between papers i and j was calculated as

$$cos(i,j) = \frac{g_{ij}}{\sqrt{g_i g_j}} \tag{2.2}$$

where g_{ij} is the number of common references papers i and j have, and g_i is the number of references paper i has. As it can be seen, this measure of similarity is not influenced by n. This implies that it is not affected by the size of the network of papers I consider. The number of references each paper has and the number of shared references are enough to describe how similar they are in terms of connectivity.

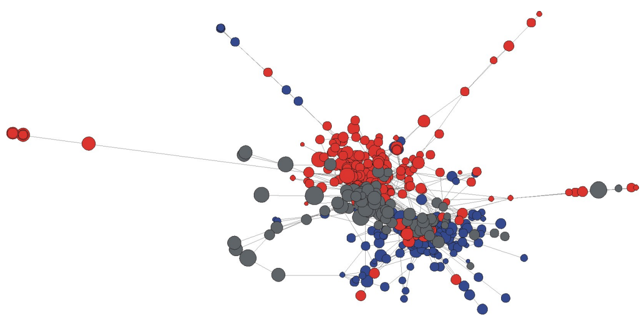


Figure 2.1: Network of academic papers colored by research department with direct distances calculated by cosine similarity -red Economics, blue Statistics, and black Business.

In Figure 2.1, the 497 papers of the data base are plotted colored by departments where they were generated. There, nodes represent papers while lines are drawn whenever it is possible to calculate a direct distance between a couple of papers. The layout in the plot is performed by a force-based algorithm which reveals computerization patterns by departments. The Business Department (grey) spreads between the Economic Department (red) and the Statistical Department (blue) suggesting that the knowledge involved in research in Business is a mix of Economics and Statistics.

Intersection

In order to capture similarity of knowledge content among papers I used an alternative measure that does not behave linearly respect to the number of shared references. Sharing g_{ij} references is significant as a proportion of the total number of references each papers cite. Thus, I propose a simple measure of similarity which considers the number of references that each paper has as a surface of a circle, and the number of shared references as the intersection's surface of the two circles. The knowledge distance between two papers, then, can be measured as the distance between the two centers of those circles as Figure 2.2 shows. This measure is very sensible for small numbers of shared references and it demands large numbers of shared references to consider two papers as very close.

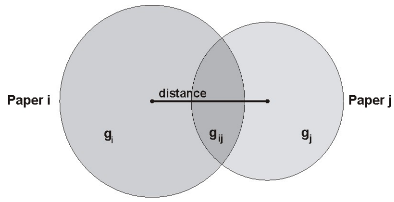


Figure 2.2: Similarity measured as the distance between the two centers of the circles

As with the cosine similarity, this measure can be calculated only for those couples of papers that share at least one references. If they do not, indirect distances are calculated using other papers that share references with both of them. Using a force-based algorithm, Figure 2.3 locates papers of the data base according to the direct and indirect similarities using this method. They are colored by departments as they were in the previous plot, and they are sized proportionally to the number of references each papers has. Similarly to the cosine similarity, Figure 2.3 shows how papers group by departments, placing Business's papers between Economic and Statistic Department.

Authors

Since the focus of this paper lies on researchers rather than on their papers, the previous analysis is used to describe scholars as located in the knowledge space. Considering papers as manifestation of knowledge, I use four different forms of studying similarities of what researchers know. The four measures are the result of combining the previous two measures of similarity, the cosine and the intersection distance, with two alternative modes of considering authors.

First, researchers are taken as the sum of their publications regarding references. All bibliographic references used by each author in their papers are summarize in one list without

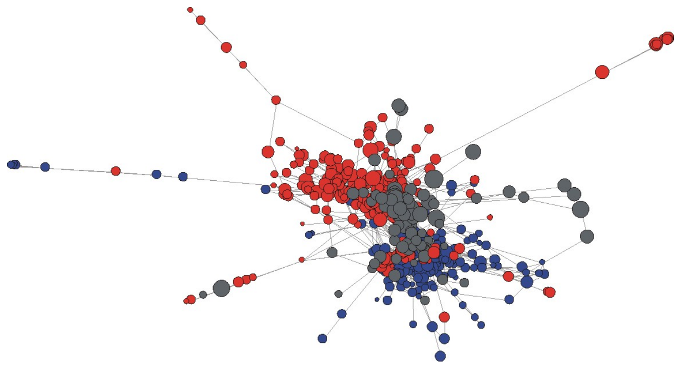


Figure 2.3: Network of academic papers colored by research department with direct distances calculated by intersections of circular surfaces and size according to total number of references.

duplication. As result, researchers' knowledge is depicted by the references they use as if the researcher would have written only one paper combining all the references she has used in many (right side of Figure 2.5). Then, shared references among authors are identified and the previous two measures of similarity are calculated for each pair of researchers.

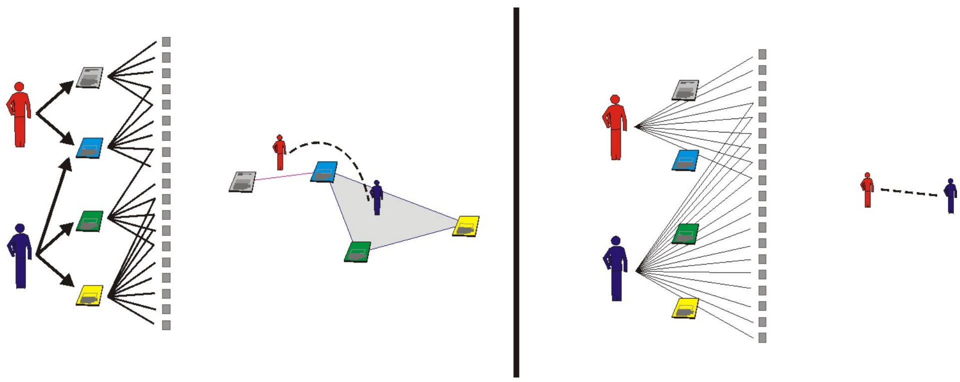


Figure 2.4: On the left side it is represented how distances are calculated among researchers as located in the centroid of their papers. In the right side distances between researchers are calculated considering all the references they use in their papers

Secondly, another approach is used to describe similarity among authors' knowledge. Instead of considering them as the sum of their papers, they are analyzed as lying at the centroid of the surface their papers form. In the previous section, similarities were calculated between papers in the data base using the cosine similarity and the intersection criterion. Based on those measures, a multidimensional scaling is performed for getting a set of coordinates for each paper that captures the calculated similarities but as an Euclidean distance. Using the generated coordinates, an author is considered to be located at the centroid of the polyhedron that her papers draw in the epistemic space as the left side of Figure 2.5 shows. Under this approach, papers are considered manifestation of researchers' knowledge and therefore, they are used to deduce a relative position towards others researchers.

Avg Degree	1.562	
Density	0.009	
Components	61.000	
Component Ratio	0.339	
Closure	0.170	
Avg Distance	3.960	
Diameter	2.000	
Compactness	0,038	
Isolated	38	(21.4%)
Isolated pairs	40	(22.5%)
Main component	47	(26.4%)

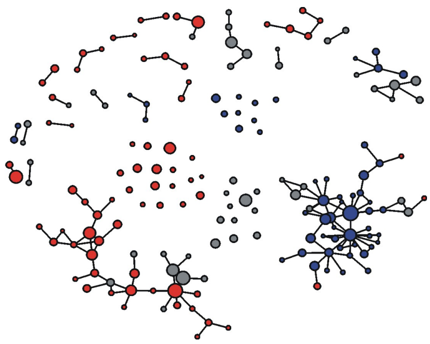


Figure 2.5: Co-authorship within the sample of researchers, colored by research departments and sized by number of references they have used in total.

Figures 2.6 plots academic researchers in the three analyzed departments considering them as the consolidation of their papers. As it can be seen, the distinction between research areas is not as clear as it is when considering papers (Figures 2.1 and 2.3). Once I have developed four alternative measures of knowledge distances among researchers, the social space must be detected in order to run an isomorphism analysis.

2.3.2 The social space

In order to capture the set of ongoing relationships among researchers in the University, I use two approaches based on two different data sources: e-mail activity and a direct survey. The first approach is applied using a data base of institutional e-mail activity among members from the three research departments -Economics, Statistics and Business- from a Spanish University during six consecutive months from September 2008 to March 2009. In order to protect privacy, e-mail addresses were anonymized with a random ID, and all content and subjects were discarded. The network was built only by considering e-mail exchange described by date/time, sender ID, receiver ID, sender's department, receiver's department, and number of receivers. Massive e-mails, i.e. those sent to more than 5 persons, were discarded assuming that collaborative activity only take place simultaneously in very small groups. E-mail accounts from administrative staff were neither considered for the sake of isolating research staff. Finally, it was only considered the internal e-mail activity within and among the three research departments. Those e-mails sent to other research departments of the same university and those sent abroad were also disregarded. As a result, a network of 396 members -150 for Economics, 97 for Statistics and 149 for Business- was built as it can see in Figure 2.7.

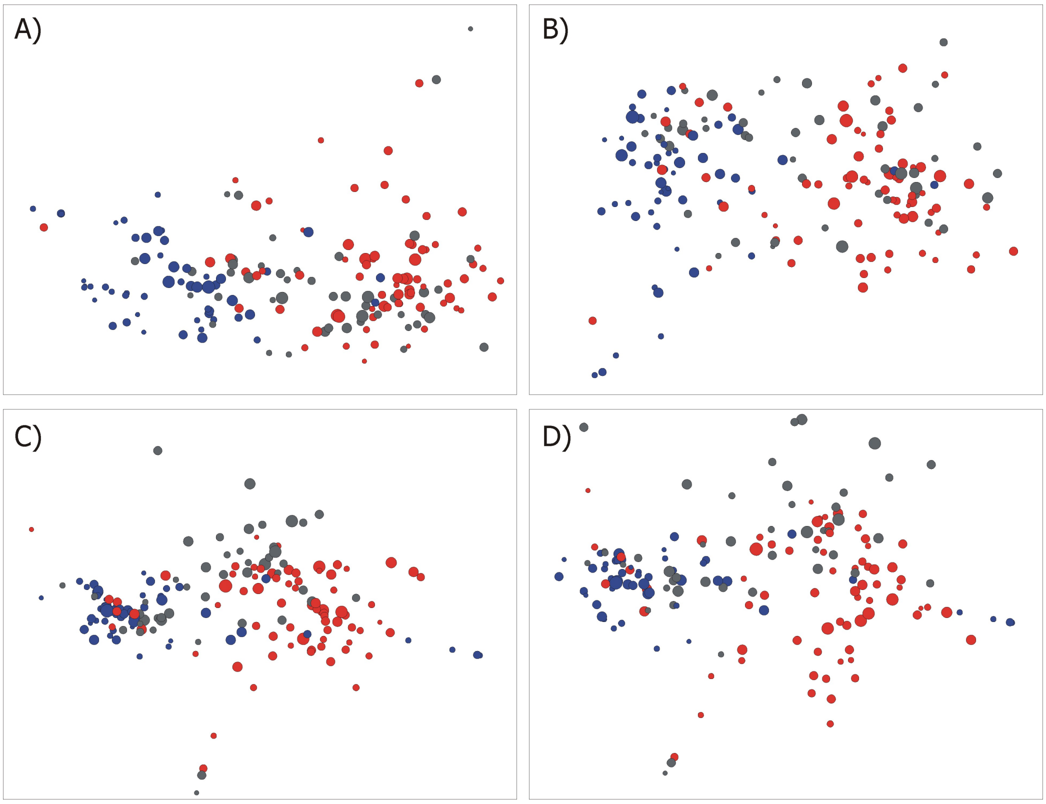


Figure 2.6: Researchers located in the knowledge space colored by department with size according to total number of references they have used in their papers and plotted by using the main 2 principal components. Panel A)uses distances calculated by cosine similarity considering researchers as the sum of their papers. Panel B) uses intersection similarity considering researchers as in panel A). Panel C) locates researchers according to the position of their papers using cosine similarity. Panel D) locates researchers according to the positions of their papers using intersection similarity

In the study of social networks, electronic records of interaction are increasingly used to capture social relations, specially e-mails networks (Perry-Smith, 2006; Kleinbaum et al., 2008; Lex et al., 2011; Guimera et al., 2006a, 2003). E-mail networks are widely used given the amount of information they capture and the relative easiness to collect the data. However, there are opposite positions in the literature about how accurate they are when describing social relations. If two persons have an intense e-mail activity it does not necessarily indicate a close social tie since it may be affected by other factor as bureaucratic routines, for example. Furthermore, two close friends might not choose e-mails to communicate but a face-to-face chat or phone calls. Thus, e-mails networks may be a poor proxy of social relations, and even if they were not, we should ask what kind of relationships are capturing.

This debate depends on the context e-mail networks are used. Particularly, in this empirical set, the e-mail network *might capture* collaborative ties. The studied research departments have almost no bureaucratic hierarchical structure. Excepting the head of the department, there is not any other designed interaction than the informal collaboration that is expected to happen

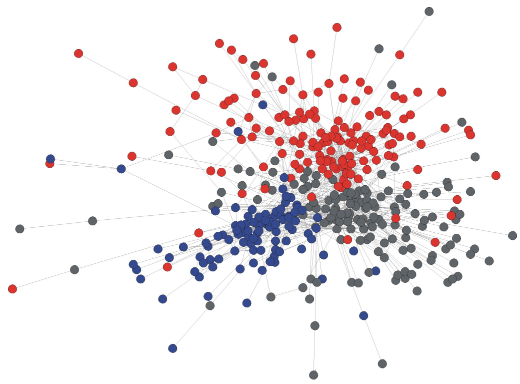


Figure 2.7: E-mail network colored by research department, being red Economics, blue Stistics, and black Business

in a research facility. It might be objected that there is a sort of bureaucratic relation between Ph.D. students and their advisers, but it can be also considered as a collaboration tie. This extremely flat organizational structure where all member do not need to periodically report tasks or exchange data on daily bases rules out bureaucratic routines that are usually captured by the e-mail network.

Another particular feature that makes this data set suitable for assessing collaborative ties is the relation between research and the use of Information and Communication Technologies (ICT). The research activity increasingly demands technological supports, specifically in those areas where these departments research in. Nowadays it is difficult to imagine an academic researcher without a personal computer in her desk. Even for those researchers who are keen on face-to-face collaboration, the use of e-mails is unavoidable when sharing papers or data that come up in the conversation that took place in aisle or at the cafeteria. That is why, in this particular research set, e-mail should be more precise than in other environments to capture informal collaboration between researchers.

However, the e-mail network fails in differentiating what kind of social ties captures. Usually, the literature works with three categories: trust-based relationships, which refers to bonds that are based on feelings; advice ties that refer to those professional relationships when someone ask for technical information related with the task she performs; and communicational ties which regards to those relationships based on sharing pertinent information about the organization they belong to but not related with the task their perform, as gossip or political issues. A semantic analysis of e-mail's contents by text mining techniques would be able to differentiate these types. However, this cannot be done for the sake of protecting privacy. For considering types of social relations, what I will assume is that *all* e-mail activity among researcher is collaborative, and some of them can be also friendship and communicational. Given that the University's e-mail service is institutional, the previous assumption seems reasonable.

As result, the social network can be interpreted as a non-Euclidean space where social distances are depicted by the whole set of relationships between the people involved. This social structure could be used for analyzing the consequences of different roles, positions or architectures of relationships in the performance of the university's researchers or departments.

Regarding the direct survey as the second approach to capture social closeness among researchers, it will be developed in the second empirical analysis.

2.4 Assessing the existence of an isomorphism

The isomorphism is studied in order to extend insights from one phenomenon to other. If two objects are isomorphic, any property that is preserved by an isomorphism that it is true for one of the objects it is also true to the other. Thus, if innovating consists in combining knowledge, modeling innovation requires a knowledge space. But his necessity can be bridged. Assuming that the social space is isomorphic to the knowledge space, researchers can model innovation using the first instead of the second.

Based on the previously proposed techniques to detect knowledge proximity between researchers, a matrix of distances between all the members of the research departments is calculated. Direct distances were obtained by applying the four different methods while indirect distances were calculated as geodesics between any pair of nodes that do not share any references. The whole set of distances depicts the knowledge space comprehended in these three research departments.

Using the e-mail network, a matrix of distances is also calculated among researchers by considering e-mail activity in a dichotomous mode. Then, given the two matrices of distances, one depicting knowledge distances and another social closeness, if both spaces were isomorphic, distances should observe very similar ordering patterns. However, with different size and without a correspondence rule between elements across sets (since the e-mail data base is anonymized) checking a possible isomorphism turns out to be logically impossible.

Even though anonymized, the list of members who are included in the e-mail data base was known. The list of members was essential to be sure that papers' authors were represented in the e-mail network. Therefore, notwithstanding the previous impossibility, it can be proved the $absence$ of an isomorphism. Should an isomorphism exist between both spaces, a subset of the social space would be isomorphic to the knowledge space. Thus, if there was not any subset of the social space isomorphic to the knowledge space, it could be stated that there is not an isomorphism between both spaces. In other words, in order to prove the absence of an isomorphism, it has to be checked the absence of an isomorphism between the knowledge space and every possible combination of the social space with size equal to the knowledge space.

The knowledge space is described by a matrix of distances between the n researchers,

$$DE = [d_e(e_i, e_j)]_{nxn} \tag{2.3}$$

where $d_e(e_i, e_j)$ stands for the epistemic distance between researchers e_i and e_j.

On the other hand, the social proximity among the m members from the three research departments is depicted by the matrix of distances,

$$DS = [d_s(s_i, s_j)]_{mxm} \tag{2.4}$$

where $d_s(s_i, s_j)$ represents the social proximity between members' e-mail accounts s_i and s_j. Since not every member have written a paper, $n < m$, i.e. the number of people in the University's email network is bigger than the number of people that have an academic paper.

Both E and S are generated by networks. Distances between nodes of these networks are calculated by the length of the shortest path (SP) between them. Since I am not interested in network characteristics but only in the distances generated by the structure, I will analyze nodes as located in a space that replicates the same pattern of distances. Thus, E is a space where researchers are located according to knowledge distances comparing their area of expertise, and S is a space where researchers are located according to their social proximity. The patterns of distances in both spaces will be study as an indirect approach to analyze a potential isomorphism between the two underlying networks that generate them.

The basic idea is that if both networks were unweighted networks, an isomorphism between them would imply the same distances among nodes. In that case, it should be easy to check an isomorphism. However, the knowledge network uses weighted links and this requires an adjustment. Given that this paper seeks to asses the assumed correspondence between social and knowledge distances and its consequences on innovation, if there is an isomorphism, it will be one that keeps the order of distances between nodes. That is:

$$\begin{aligned} &\text{if } d_s(s_i, s_j) \leq d_s(s_h, s_k)], \\ &\text{then } d_e(e_i, e_j) \leq d_e(e_h, e_k)] \\ &\text{for } \forall i, j, h, k \end{aligned} \tag{2.5}$$

In other words, given an isomorphism between two networks, if two persons are closer in their area of research than other two colleagues, then, the first couple should be socially closer than the second one.

However, to check the existence of an isomorphism of this type between the two involved networks represent two big challenges. In the first place, one of the networks is anonymous and this implies a large permutational analysis. Secondly, the set of nodes from the e-mail network is larger than the epistemic network's one ($n < m$). This, in turn, represent the biggest challenge: a combinatorial number of $\binom{m}{n}$ possibilities.

These two characteristics of the data set make logically impossible to prove the existence of an isomorphism, but, as it was said before, it can be proved its absence. The next section explains the approach.

2.4.1 Algorithm

Analyzing a possible isomorphism between the e-mail network and the knowledge network has two major obstacles. First, the two networks have different size. Second, the lack of a correspondence rule between their nodes because of the anonymity. That is why, in order to tackle these problems, I will assume that E and S are isomorphic and then, I will explore the necessary consequences. If they do not apply to the data set, the hypothesis will be disregarded and, by $modus\ tollens$, the opposite statement will be taken as true; i.e. the networks *are not* isomorphic.

Given that there is not a correspondence rule between nodes of one network and the other, it is not possible neither to match nodes nor to check whether the relation between distances is preserved. Thus, I work by process of elimination discarding possible correspondences between pair of nodes from both networks. Specifically, taking the pair $e_I \in E$ and $s_J \in S$, if e_I and s_J are the same node (represented in both networks), and S and E are isomorphic, it should be a subset $S_I \subset S$ of n members including s_J that shows the same pattern of distances among its members than the pattern of distances among E's members. In order to do the latter, I postulate not only that e_I represents the same researcher in E than s_J in S, but also that the subset S_I represents the same researchers of E. However, since I do not know which node of S_I is each node of E, I cannot perform a direct comparison between distances. That is why, I will compare the pattern of distances, i.e., how nodes in S_I are distributed according to distances to s_J; and how nodes in E are distributed according to e_I. The basic idea is that if S_I and E were isomorphic, the closest node to s_J in S_I should correspond to the closest node to e_I in E; as well as the second closest node to s_J in S should correspond to the second closest node to e_I in E, and so on. Ignoring measures of distances but considering only the order of proximity, if S_I and E were isomorphic, and nodes correspond to each other according to the proximity to e_I and s_J, it should also be observed that distances among nodes keep the order as well.

Figure 2.8 portraits an example. There, if the second closest node to s_J in S_I is further to the third closest node to s_J than to s_J, while in E, the second closest node to e_I is closer to the third closest node to e_I than to e_I, then, neither E and S_I are not isomorphic. Consequently, e_I cannot correspond to s_J. Assuming isomorphism, the pair e_I and s_J is discarded as being the same researcher in the social and epistemic network, and another pair is taken to conduit the same test. The test is performed for every possible pair of E's members and S's members.

Formally, for $e_I \in E$ and $s_j \in S$, I create the vectors Q_I and Q_J

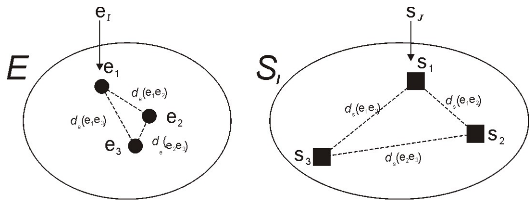


Figure 2.8: Correspondence of e_I to s_J.

$$Q_I = [(e_k)_h]_{1xn} = [q_h] \text{ such that} \begin{cases} e_k \in E \\ d_e(e_I, q_h) < d_e(e_I, q_{h'}) \quad \text{if} \quad h < h' \end{cases} \tag{2.6}$$

$$P_J = [(s_k)_h]_{1xm} = [p_h] \text{ such that} \begin{cases} s_k \in S \\ d_s(s_J, p_h) < d_s(s_J, p'_h) \quad \text{if} \quad h < h' \end{cases} \tag{2.7}$$

Naturally, the first element of Q_I (q_1), is e_I since e_I is the closest node to itself. The same happens with first element of Q_J, i.e. $p_1 = s_J$.

Since $m > n$, in order to compare vectors of equal length I have to sequentially take different subsets of size n from Q_J's members. Then, for avoiding taking each of the $\binom{m}{n}$ possible subsets, I will follow another strategy. If E and S were isomorphic, and e_I and s_J were the same node in the two networks, I should be able to find the first g nodes in Q_I, in Q_J, although not necessarily consecutively. Thus, I take every possible ordered subset of Q_J of size g. Calling this subset as

$$G_{1xg} = [(p_h)_l] = [g_l] \text{ such that} d_s(s_J, g_l) < d_s(s_J, g'_l) \text{ if} l < l'. \tag{2.8}$$

Under these assumptions, and as it was explained before, it should be observed that the set of $g(g-1)/2$ distances among members of G can be sorted by length equally to the set of distances from the first g elements of Q_I. For doing the latter, the matrices of distances of both subsets must be compared. Being,

$$D_G = [d_s(g_i, g_j)]_{gxg} \tag{2.9}$$

D_G is converted into the sorted vector

$$[d_G]_{1xg(g-1)/2} = [(d_s(g_i, g_j))_r] \begin{cases} 1 \leq i \leq (g-1) \\ i+1 \leq j \leq g \\ 1 < r < g(g-1)/2 \\ (d_s(g_i, g_j))_r < (d_s(g'_i, g'_j))'_r \quad \text{if} \quad r < r' \end{cases} \tag{2.10}$$

Doing the same for the subset of the first g elements of Q_I, I get the vector

$$[d_{Q(g)}]_{1xg(g-1)/2} = [(d_e((q_i, q_j))_r] \begin{cases} 1 \leq i \leq (g-1) \\ i+1 \leq j \leq g \\ 1 < r < g(g-1)/2 \\ (d_e(q_i, q_j))_r < d_e(q'_i, q'_j))'_r \quad \text{if} \quad r < r' \end{cases} \tag{2.11}$$

Finally, comparing d_G with $d_{Q(g)}$, if E and S were isomorphic, and e_I and s_J were the same node represented in E and S, it should be observed that, for elements in the same position in both vector (same r), sub indexes i and j should coincide. In other words, the distance between elements i and j of both vectors should keep the same relation (larger or smaller) to the rest of distances in the set.

If I cannot find $g < n$ nodes that, as a subset, does not show the same order of distances than the first g nodes of Q_I, it does not make sense to continue searching for larger sets that accomplish this. Otherwise, if I do find subsets of g nodes that keeps the same ordering of distances than Q_I's elements, I enlarge their size to $g+1$ and look again. Finally, if I get at least one subset of S of size $g = n$ that do not brake this consequence of being isomorphic, I cannot neglect the assumption that e_I and s_J are the same node in E and S. On the contrary, I reject the correspondence of e_I and s_J.

Under the assumption that E and S are isomorphic, after this process of elimination, it should remain at least one possible correspondence between nodes in E and S. Since $n < m$, all nodes of E should have a candidate for being the correspondent node in S, but not all nodes in S should have one in E. Therefore, if this is not observed in the test, the hypothesis of isomorphism between E and S is discarded.

2.4.2 Empirical results

Following the previous methodology, the data set is divided by research departments in order to reduce the number of possibilities to consider. The Statistics Department, for instance, has 56 researchers that are represented within the 97 anonymous email accounts. As result, each of the 56 researchers could correspond to each of 97 e-mails accounts, a total of 5,432 possible correspondences to check. As explained before, every possible pair is tested for a potential local isomorphism, increasing the size of the subset until the isomorphism is not hold anymore. Registering the maximum size of a possible correspondence between a node in the knowledge network and a node in the e-mail network, Table 2.4 shows the obtained results. Due to the time this process consumes, only the Statistical Department was tested using the four alternative approaches, while the other two departments were only tested by using two metrics.

Table 2.1: Empirical results

Metric/department	Median	Mean (%)	Std. Dev.	Max. (%)	Size KN	Size SN	Cases
cosine similarity (s)	10	10.55 (18.8%)	2.01	15 (26.8%)	56	97	5,432
intersection (s)	10	10.54 (18.8%)	1.99	14 (25.0%)	56	97	5,432
cosine surf. (s)	11	10.86 (19.4%)	1.67	14 (25.0%)	56	97	5,432
intersection surf. (s)	11	10.88 (19.4%)	1.67	14 (25.0%)	56	97	5,432
cosine similarity (e)	11	12.49 (17.1%)	3.89	21 (28.8%)	73	150	10,950
intersection (e)	12	13.29 (18.2%)	3.47	20 (27.4%)	73	150	10,950
cosine similarity (b)	12	12.11 (24.7%)	1.93	15 (30.6%)	49	149	7,301
intersection (b)	12	12.09 (24.7%)	2.03	16 (32.7%)	49	149	7,301

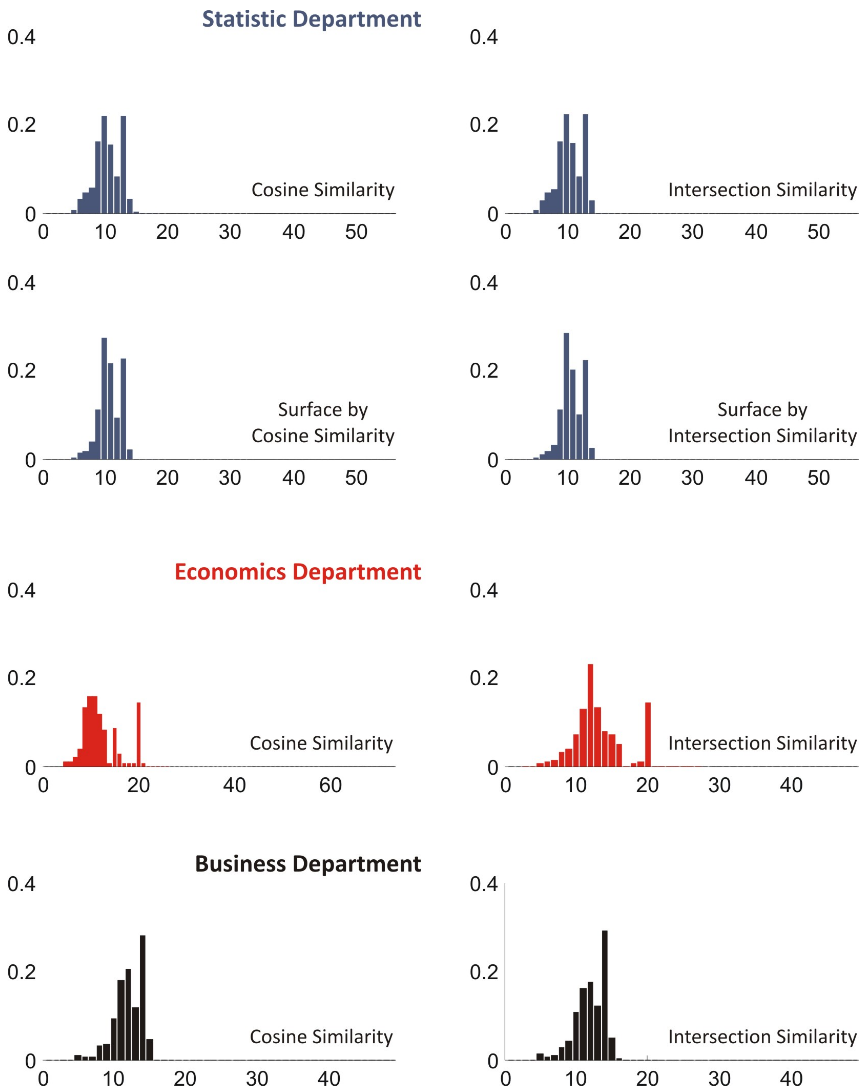


Figure 2.9: Maximum number of nodes in a subset when checking for a possible isomorphism before it is discarded, for every pair of nodes in both spaces, and using the four different measures of knowledge proximity.

In order to discover a possible isomorphism, as it was explained in Section 2.4.1, every node that compose the knowledge network should find a least one node from the email network that shows the same order of distances with a set of the same size withing this network. For instance, each of the 56 nodes that describe the knowledge space within the Statistic Department should find at least one node in the social network that, along with 55 nodes more, reproduce the same order of distances. As mentioned before, for this department, the four alternative measures for capturing similarity of knowledge among researchers were considered. Table 2.4 shows results. For each of the 56 members of the knowledge space, the algorithm tested a possible correspondence with each of the 97 nodes in the social space. A total of 5,432 cases were considered. For each possible correspondence, the algorithm tested how long the isomorphism held when increasing the members considered. As it can be observed, 15 and 14 were the maximum number achieved, 25% of the total size. Since the algorithm was designed to discard an isomorphism, in order not to discard it (which it does not imply its confirmation), at least one isomorphic subset of 56 members of the e-mail network should be found for each of the 56 nodes of the knowledge network. Clearly, this is not the case given the absence of any isomorphic subset. It can be concluded that the social space of the Statistic department is not isomorphic to its knowledge space conditioned to the mode they were constructed.

The absence of an isomorphism could be attributed to the way I calculated the social and knowledge proximity between researchers. However, the result is still valid since the e-mail network of the department could be used for explaining the performance of scholars of this University as a accepted methodology. What this result shows is that it fails in capturing knowledge expertise and potential compatibility with other members of the department according to another broadly accepted technique.

If discarded the isomorphism between the social and the knowledge space within the Statistical Department, logically, it can be discarded the isomorphism of the whole data set (i.e. the three research departments considered simultaneously). Notwithstanding rejecting the hypothesis of isomorphism of one of its subsets allows rejecting it for the whole, the same analysis is performed for the other two research departments. Table 2.4 shows the results. As it can be observed, results do not differ much. The possibility of a possible isomorphism is strongly rejected.

The results depicted in Table 2.4 may also suggest that *if* there is an isomorphism, it only holds locally. This idea cannot be strongly stated because of the anonymity of the e-mail network. However, the short range of possible isomorphism may indicate that they may exist in small sub-sets.

Due to the nature of the data, in order to discard the possibility of miss-capturing the social network by using the e-mail network, I also analyzed the possibility of an isomorphism by also conducting a sociometric survey as the following Section describes.

2.5 Another approach: direct survey

The anonymity of the e-mail data base makes impossible a deeper analysis in the comparison of both the knowledge and the social structure. It represented a big computational challenge for an very important but simple result. In order to enrich this analysis, I used a complementary information obtained from a survey.

In the sociological literature, ego-network sociometric surveys has a long record as an instrument to capture social networks (Ferligoj and Hlebec, 1999; Granovetter, 1983; Sampson, 1988; Podolny and Baron, 1997; Morrison, 2002). Standard procedures consist in asking a person about different kinds of relationships she or he has with others. Usually, in order to make the process easier, possible answers for the questions are listed. Then, if a person is asked who is her or his friend, she will pick up the names among the entire list of candidates[1]. For this reason sociometric surveys can be very time-consuming for the respondent and therefore it can low the response rate. This may represent a big problem to the statistical analysis since in network analysis the omission of only one edge can radically alter the whole structure.

In order to avoid the problems associated with sociometric surveys but capitalize its benefits, I conducted an ego-network sociometric survey but with a different approach. This approach is designed not only to be less time consuming but also statistically invulnerable to low response rates. Based on the four alternative knowledge distances proposed at Section 2.3.1, this method consists in selecting the closest people to each member in their research topics. Then, each member is asked what kind of social relation she or he has with them. Figure 2.10 illustrates this idea. There, researchers are plotted using the two main principal components so their euclidean distances capture their knowledge distances. Then, for each researcher, the ten closest colleagues are detected and summarize into a personal survey where she or he is asked about the kind of relationships she or he has with them. Instead of listing the entire set of possible researchers they might have a relation with (178 persons), this list only picks the most proximate people according to their area of expertise. If the isomorphism holds, it should be observe that people has close collaborative ties with those who are proximate in the knowledge space. Furthermore, given that the answers of researchers are not used to build the social network but only to check if the isomorphism holds locally, low response rates will not undermine the validity of the analysis.

Therefore, for each researcher of the data base, I generated a list of names containing the 10 closest colleagues in the knowledge space according to the four metrics: a total of 40 names. Of course, some of them were repeated since the four criteria can coincide when choosing the closest neighbors. Thus, the list was reduced consolidating repetitions and sorting the names according to i) how many criteria have chosen a name as proximate, or, if this criterion does not differentiate, ii) to the closest members according to an average standardized measure of proximity. The resulting personal questionnaires do not include more than 15 to 20 names.

[1]It is called roster recalled method (Wasserman and Faust, 1994)

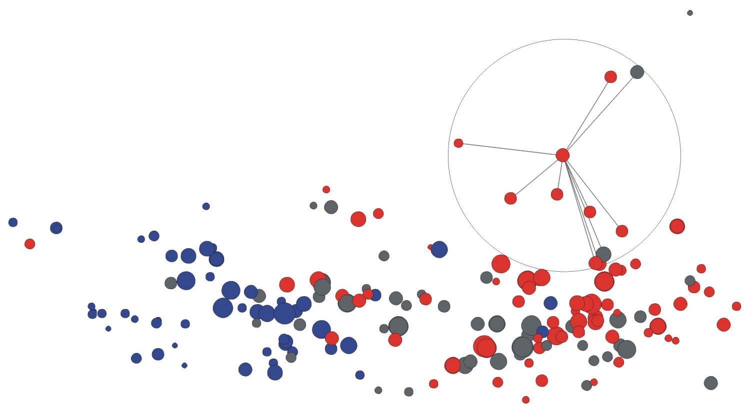


Figure 2.10: Selecting the 10 closest researchers to certain node according to their coordinates in a knowledge space

The survey contained three questions regarding the listed people: i) how often the researcher turns to those colleagues for research-related consultation or discussion; ii) how often researcher talks to those colleagues about university-related questions different from research; and iii) how often the researcher converses with them about personal life. The first question is intended to capture the collaborative relationships based on technical advice and academic knowledge. The second one tries to capture information-based relationships where people exchange information related to the institution they belong to. The third question seeks to reveal trust-based relationships that exceed their occupation. Surveyed researchers were invited to answer using a Likert scale of 4 grades: never, seldom, sometimes and often.

2.5.1 Empirical analysis and results

With a rate of response of 64.04% (114 answers out of 178) on two waves of sent surveys, the collected data is plotted in Figure 2.11. As explained before, social interaction is a discrete variable graded in 4 levels while the measure of knowledge distance is a continuous variable. In order to have a first look, knowledge distances is graded also in 4 levels in Figure 2.11 by mimic the distribution of social interaction. For example, an interviewed person declared that out of the 10 closest people in terms of knowledge, she *often* discusses academic ideas with 2 of them (category 1), *sometimes* she turn to 3 of them for research advice (category 2), while she *never* talks about academic topics with the rest (category 4). Then, knowledge distances are ordered in order to have the same three segments, where the closest 2 persons are classified as 1, the following 3 as 2, and the remaining 5 classified as 4. If the isomorphism holds, those colleagues classified in terms of social distances in i^{th} category should be also classified in the i^{th} category of the knowledge distances.

Knowledge distances were also categorized in 4 levels observing the order of proximity given by the correspondent measure. Figure 2.11 shows the results in 12 Panels combining the

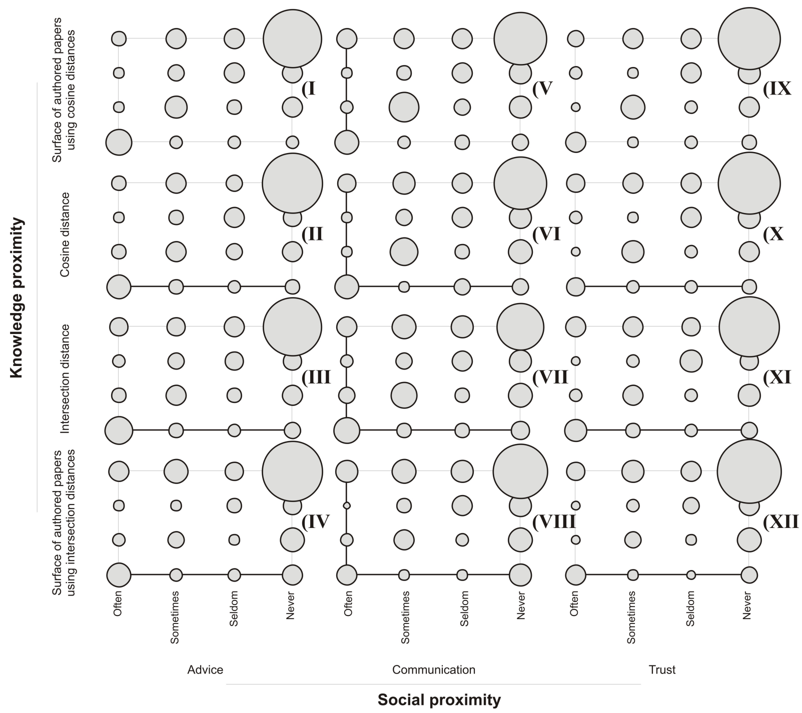


Figure 2.11: Contrast between the results of the conducted survey and knowledge proximity of researchers

four measures of knowledge proximity and the three kinds of social relationships researchers may have. The size of the circles represents the relative density of observations located in each intersection. Visually, it can be noticed a lack of correspondence between orderings suggested by the heavy presence of observations outside the main diagonal.

Table 2.2: Descriptive statistics

	Variable	Obs	Mean	Std. Dev.	Min	Max
(1)	Cosine distance	926	31.2568	9.3468	0	71.4384
(2)	Intersection distance	926	35.8353	11.0712	0	82.6864
(3)	Cosine surface	926	12.0745	4.7161	2	36
(4)	Intersection surface	926	25.5928	11.4664	0	85
(5)	Principal component 1	926	0.0000	15.5828	-52.0040	61.5864
(6)	Principal component 2	926	0.0000	10.3460	-33.4175	37.8895
(7)	Advice	926	2.5745	1.1162	1	4
(8)	Communication	926	2.2495	0.9412	1	4
(9)	Trust	926	2.7387	1.1027	1	4

In order to empirically assess how similar the social and the knowledge space are in terms of distances among their elements, I use an ordered logit model. This model allows dealing

Table 2.3: Linear correlation among variables

	Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1)	Cosine distance	1.0000						
(2)	Intersection distance	0.9357*	1.0000					
(3)	Cosine surface	0.4462*	0.3991*	1.0000				
(4)	Intersection	0.3129*	0.3463*	0.7825*	1.0000			
(5)	Principal component 1	0.8966*	0.9149*	0.6760*	0.6746*	1.0000		
(6)	Principal component 2	-0.3921*	-0.3738*	0.4957*	0.7349*	0.0000	1.0000	
(7)	Advice	0.0743*	0.0751*	0.1155*	0.1545*	0.1222*	0.0954*	1.0000
(8)	Communication	0.0066	-0.0039	0.0375	0.0293	0.0133	0.0298	0.6054*
(9)	Trust	0.0221	0.0263	0.0499	0.0843*	0.0532	0.0629	0.6667*
	Variable	(8)	(9)					
(8)	Communication	1.0000						
(9)	Trust	0.6483*	1.0000					

* p<0.05

with an ordinal discrete variable as the dependent one. The probability of observation i of being classified in category c is a linear function of independents variables plus a random error (Williams, 2006). In this particular case, the only independent variable will be the knowledge distance measured in each of the four alternatives that were proposed before. Furthermore, a new measure of knowledge distance is proposed by considering simultaneously the former four measures in their two main principal components which capture the 96.21% of variation. Table 2.2 describe the four knowledge measures plus the two principal components of those four along with the three kind of social interaction registered in the survey. Pairwise linear correlations among these variables are shown in Table 2.3. As it can be noticed, the four measures of knowledge distance are significantly and positively correlated, as well as the three measures of social proximity. Correlations among these two blocks are negligible.

For each of the three types of social relationships (academic related, institution related and friendship), five models were estimated according to each of the knowledge distances and their two principal components. Table 2.4 shows estimated coefficients and the three cutpoints that determine the probability of being classified in one of the four categories of social proximity. As it can be observed, all measures of knowledge distance are significantly and positively related with academic related social relations. In the survey, researchers were asked "*how often do you turn to these people for research-related consultation or discussion?*". This question is expected to capture social interaction based on technical knowledge of the counterpart. Results of the regressions indicate that social distances in this dimension are positively and significantly correlated with knowledge distances.

The significant relation between knowledge similarity and social proximity among researchers does not hold for the other two types of social interaction captured in the survey. The second column of Table 2.4 displays results with respect to social interaction where its main content is information related with the University researchers work for. Survey researchers were asked "*How often do you talk with these people about university-related questions different from research?*".

Table 2.4: Parameter estimates of different metrics for knowledge distance on different types of social proximity. Ordered logit regression

		Dependent variable: Social Distance (4 levels)		
		Advice	Communication	Trust
Knowledge Distance	**Cosine distance**	**0.0149 (0.006) ***	**0.0019 (0.006)**	**0.0048 (0.006)**
	cut1 (95% Conf. Interval)	-0.768 (0.210 -1.181)	-1.029 (0.211 -1.443)	-1.379 (0.213 -1.796)
	cut2 (95% Conf. Interval)	0.349 (0.208 -0.059)	0.450 (0.209 0.041)	-0.190 (0.206 -0.594)
	cut3 (95% Conf. Interval)	1.447 (0.214 1.027)	2.290 (0.227 1.844)	0.847 (0.208 0.439)
	(Prob > chi2 | Log likelihood)	(0.0194 | -1278.8872)	(0.7582 | -1206.8627)	(0.4413 | -1261.0513)
	Intersection distance	**0.0128 (0.005) ***	**0 (0.005)**	**0.0046 (0.005)**
	cut1 (95% Conf. Interval)	-0.775 (0.205 -1.177)	-1.090 (0.205 -1.492)	-1.365 (0.207 -1.770)
	cut2 (95% Conf. Interval)	0.343 (0.203 -0.054)	0.389 (0.202 -0.007)	-0.176 (0.200 -0.569)
	cut3 (95% Conf. Interval)	1.441 (0.209 1.032)	2.229 (0.221 1.796)	0.861 (0.202 0.465)
	(Prob > chi2 | Log likelihood)	(0.0176 | -1278.8025)	(0.9995 | -1206.9101)	(0.3853 | -1260.9709)
	Cosine surface	**0.0468 (0.013) *****	**0.0149 (0.013)**	**0.0203 (0.013)**
	cut1 (95% Conf. Interval)	-0.678 (0.169 -1.010)	-0.912 (0.170 -1.246)	-1.288 (0.173 -1.628)
	cut2 (95% Conf. Interval)	0.447 (0.168 0.119)	0.569 (0.168 0.239)	-0.097 (0.166 -0.423)
	cut3 (95% Conf. Interval)	1.552 (0.175 1.208)	2.410 (0.192 2.034)	0.943 (0.170 0.610)
	(Prob > chi2 | Log likelihood)	(0.0003 | -1274.9919)	(0.2447 | -1206.2334)	(0.1107 | -1260.0756)
	Intersection surface	**0.0261 (0.005) *****	**0.005 (0.005)**	**0.0132 (0.005) ***
	cut1 (95% Conf. Interval)	-0.582 (0.154 -0.883)	-0.962 (0.154 -1.264)	-1.198 (0.156 -1.504)
	cut2 (95% Conf. Interval)	0.552 (0.152 0.253)	0.518 (0.151 0.222)	-0.003 (0.149 -0.296)
	cut3 (95% Conf. Interval)	1.665 (0.161 1.349)	2.358 (0.176 2.013)	1.040 (0.154 0.739)
	(Prob > chi2 | Log likelihood)	(0.0000 | -1269.7092)	(0.3416 | -1206.4578)	(0.0115 | -1258.1544)
	Principal component 1	**0.0155 (0.004) *****	**0.002 (0.004)**	**0.0064 (0.004) +**
	Principal component 2	**0.0176 (0.006) ****	**0.0051 (0.006)**	**0.0106 (0.006) +**
	cut1 (95% Conf. Interval)	-1.253 (0.079 -1.408)	-1.091 (0.076 -1.239)	-1.535 (0.086 -1.704)
	cut2 (95% Conf. Interval)	-0.119 (0.066 -0.249)	0.390 (0.067 0.259)	-0.340 (0.067 -0.471)
	cut3 (95% Conf. Interval)	0.995 (0.074 0.849)	2.231 (0.111 2.013)	0.702 (0.070 0.565)
	(Prob > chi2 | Log likelihood)	(0.0000 | -1269.5201)	(0.5928 | -1206.3872)	(0.0465 | -1258.2803)
				(Standard Error) *** p<0.001 ** p<0.01 *p<0.05 +p<0.1

The third column shows results regarding more intimate social relationships. For capturing them, researchers were asked *"How often do you converse with these people about personal life?"*. As it can be observed, although all of estimated coefficients are positive, none of them is significant with only two exceptions. One of them is the relation between friendship relationships with knowledge distances as captured by the two main principal components of the four measures. The other one is the relation between friendship and knowledge distance captured by the intersection surface. However, although its statistical significance, those coefficients are considerably smaller than the ones estimated for academic-related social interactions. The rest of the estimates seemingly describe the absence of a significant relation between knowledge distances and social distances in these two dimensions.

Figure 2.12 helps interpreting these results. Among the four alternative measures for capturing knowledge distances, and their two principal components, I choose to use the model estimated for the principal components given that they are a linear combination of the previous

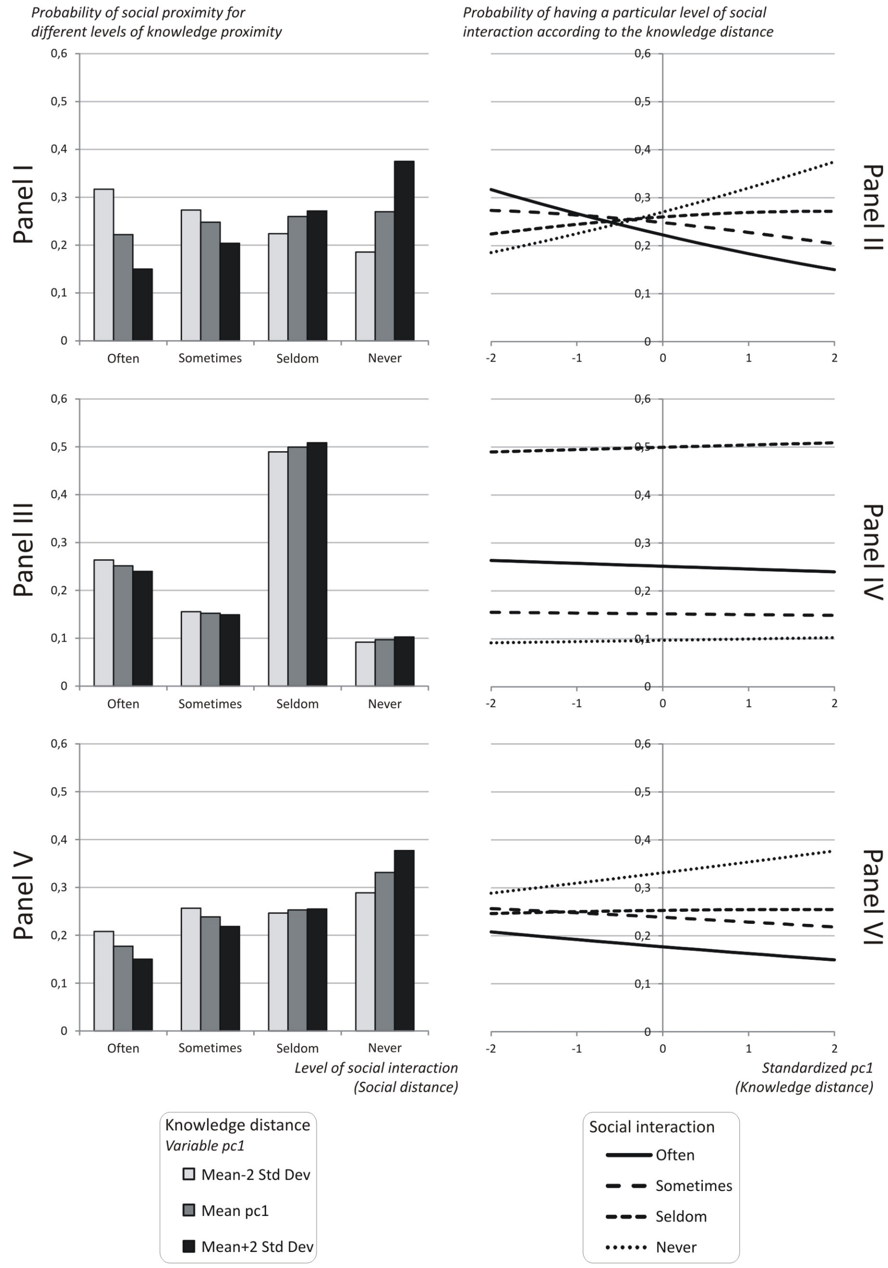


Figure 2.12: Graphical interpretation of results

four that maximizes the variance. As Table 2.3 shows, this measure is highly correlated with all the knowledge metrics. Since its interpretation in terms of magnitudes only reflects relative distances as the rest of the metrics, in the following paragraphs, knowledge distance is considered as the first principal component. Figure 2.12 is composed by 6 Panels. Each of the three row plots empirical results for each of the three types of social relations, being the first related with academic-related interactions, the second with institution-related interacions, and the third with friendship relations.

The column on the left graphs the expected probability of classifying a social relation in each of the four possible categories when the knowledge distance is equal to the mean of the observations, the mean plus two times the standard variation, and the mean minus two times the standard variation. For instance, in Panel I the black bars describe the probability of a high knowledge distance of being classified in each of the categories. Researchers who are distant in the knowledge space, in this model, are expected with more than twice the probability to generate distant academic-based social relationships than close academic-based social relationships. Conversely, when two researchers are close in the knowledge space, the probability of having a close academic-based social relation is significantly higher than a distant relation. In other words, this model supports the existence of an isomorphism between the knowledge space and social space generated by professional advice networks. Panel III and Panel IV do not show this pattern. As general rule, in order to observe a possible isomorphism between the two spaces, black bars are expected to achieve the maximum altitude in categories "*Seldom*" or "*Never*", while light grey bars are expected to maximize altitude in categories "*Often*" and "*Sometimes*". Gray bars, the ones capturing the mean knowledge distance, are expected to achieve maximum values around the center, i.e. in categories "*Sometimes*" and "*Seldom*". The left column of plots shows that this pattern only applies to Panel I. Panel III and Panel IV brake the expected pattern indicating the lack of evidence for supporting the possible isomorphism between the two spaces.

The column on the right in Figure 2.12 displays the same results but using another perspective. At each Panel, the horizontal axis describes the variation of the knowledge distance from -2 standard deviations to 2 deviations. Four lines depict as a function the expected probability of being classified in each of the four categories of social proximity according to the knowledge distance of researchers. Panel II, for instance, shows the decay of the probability of being classified in "*Often*" when the knowledge distance increases. This indicates that low knowledge distances are associated with close social relationships for academic advice. The rest of the lines describe an expected behavior if the two spaces were isomorphic. By contrasting the forms of this estimated functions with the ones at Panel IV and Panel VI, it can be observed the clear lack of correspondence between the social space and the knowledge space regarding these kind of social relationships. Although academic-related social proximity is the only one that depicts the pattern expected for isomorphic spaces, the estimated probability functions are far from being perfect. If the two spaces were isomorphic, the line of "*Often*" (close social relationships) would achieve values near 1 for low values of the knowledge distance and negligible values for the rest. The lines corresponding to "*Sometimes*" and "*Seldom*" would describe an inverted U shape with maximum values near the mean, being those values close to 1. As it can be easily observed, this is not the case for any of relationships between social distances and knowledge distances.

A final comment must be included in the empirical assessment of the isomorphism between the social and knowledge space. The co-authorship network was not used for this analysis mainly because its low density (0.9% - Figure 2.5). This not only impedes calculating enough distances for a robust statistical regression, but also impedes capturing social relationships. Co-authorship implies a strong collaborative tie between researchers and can be used for describing the social space. However, if papers produced by more than one researcher are used for capturing what they know, it is expected a high correlation between co-authorship and knowledge expertise. This correlation would not necessarily be significant in larger volumes of data where researchers could differentiate to their co-authors working with other researchers or alone. This is not the case of this chapter. The the data set used for the empirical analysis considers only three years of research. Therefore, the co-authorship tie cannot be considered independently to the knowledge measure of similarity.

Table 2.5

Dependent variable: social distance measured by the co-authorship network

	Knowledge Distance				
	Principal components	Cosine distance	Intersection distance	Cosine surface	Intersection surface
cons	93.1812 (0.635) ***	90.6014 (1.172) ***	90.2963 (1.158) ***	88.0744 (0.993) ***	89.4093 (0.95) ***
var_1	0.099 (0.018) ***	0.089 (0.031) **	0.085 (0.025) **	0.410 (0.056) ***	0.151 (0.023) ***
var_2	0.101 (0.025) ***				
Dummy for Technical Advice					
2	2.010 (0.979) *	2.184 (1.001) *	2.211 (1.002) *	2.161 (0.982) *	1.950 (0.989) *
3	2.022 (1.104) +	2.391 (1.115) *	2.405 (1.112) *	2.112 (1.11) +	2.029 (1.11) +
4	2.878 (1.188) *	3.611 (1.216) **	3.588 (1.211) **	3.154 (1.195) **	2.937 (1.188) *
Dummy for Communication					
2	-1.178 (0.904)	-1.454 (0.913)	-1.466 (0.913)	-1.256 (0.899)	-1.163 (0.915)
3	-2.362 (1.02) *	-2.815 (1.036) **	-2.782 (1.036) **	-2.590 (1.022) *	-2.333 (1.028) *
4	-1.906 (1.306)	-2.269 (1.337) +	-2.202 (1.333) +	-2.282 (1.304) +	-1.931 (1.312)
Dummy for Trust					
2	0.748 (0.962)	0.804 (0.984)	0.788 (0.984)	0.898 (0.947)	0.701 (0.973)
3	0.709 (1.090)	0.986 (1.118)	0.957 (1.119)	0.998 (1.096)	0.604 (1.095)
4	1.383 (1.001)	1.386 (1.029)	1.368 (1.027)	1.580 (0.997)	1.304 (1.007)
Dummy for direct or indirect connection					
connec	-72.567 (0.861) ***	-73.608 (0.821) ***	-73.487 (0.84) ***	-72.591 (0.838) ***	-72.727 (0.861) ***
obs	926	926	926	926	926
R2	0.9517	0.9499	0.9500	0.9519	0.9514
(Standard Error) *** p<0.001 ** p<0.01 *p<0.05 +p<0.1					

Table 2.5 empirically shows how distances in the co-authorship network are related with social distances and knowledge distances. Five models are run for co-authorship distance as dependent variable and changing the main explanatory variable that captures knowledge distances. On the left column, var_1 represents the independent variable that stands for knowledge distance and changes in each model (column) for all the alternatives. var_1 is only used for the second principal component. Three dummy variables are included for each model introducing social distances as captured by the survey. Finally, an extra dummy variable distinguish

whether two researchers are connected through the network (either directly or indirectly). Since the co-authorship network is composed by several unconnected blocks (Figure 2.5), geodesic distances among them cannot be calculated. Instead, they are assumed to be equal to the maximum geodesic distance calculated. For this reason, the former dummy treats differently those that belong to the same block than those who do not. As it can be observed, co-authorship is strongly related with knowledge distances across all models. The technical-advice relation is positively and significantly related with co-authorship distances, while the other two fail in being positive or significant. The relations estimated in this regression simply show that the co-authorship network is related with knowledge distances and social distances considering those based on research matters.

2.6 Discussion

This chapter explores how far can be assumed an isomorphism between the social space and the knowledge space when analyzing knowledge generation. The increasing use of social network analysis for innovation makes this question a necessary one. As explained in the introduction, social networks offer the possibility of observing the very process of knowledge combination for generating new one. By simply assuming that the complexity of knowledge necessary for innovating easily overwhelm a single person and therefore, she or he has to interact with other in order to be able to collectively process it, the social structure offers a way of tracking combinations. Naturally, this approach is highly valuable for studying innovation as the large amount of research on this field proves it. However, as reviewed in previous sections, the study of innovations using social structures may have overlooked the underlying assumption that drives all the logic involved in the analysis.

This chapter tested four different ways of capturing social relationships and their possible correspondence with knowledge similarity. Section 2.3 uses the institutional e-mail network while Section 2.5 uses a survey differentiating three kind of relationships: professional (technical advice), organization-related (communication), and friendship (trust). All four approaches to capture social interaction could be used for analyzing innovation. However, as empirical results show, most of them fail in explaining the similarity of knowledge of researchers. The only exception is the proximity captured by the survey regarding technical-advice relations.

These results suggest a very intuitive idea. Distances in a social space are correlated with those in the knowledge space as long the social space is defined on relationships that captured exchange of information related with knowledge. Consequently, a social network of this kind of relationships would mimic the knowledge space and therefore it could be used for analyzing innovation. However, the methodology used in this paper cannot fully support the previous idea. Social networks structure information of ego networks into indirect relations among components. In fact, indirect relations are expected to compose the great majority of relations. The survey that collected data about frequency of relationships between researchers cannot be combined into a network structure that allow capturing indirect relations. For instance, this methodology states that if researcher A never consults researcher B, then, they are distantly located in the social space. However, it may happen that both A and B frequently consult a third researcher C. The short distance between A and C and B and C would imply an indirect proximity between A and B. Since this cannot be captured, the empirical analysis is based on direct relations, i.e. local information. This data set supports the existence of an isomorphism, at least, locally.

In order to go beyond the locality of the information provided by the survey, the e-mail activity was used to sketch the whole network of social relationships among the three research departments. The anonimity of email accounts impeded a simple statistical analysis although this problem was addressed using a different approach giving coherent results. The algorithm, by ordering distances in increasing diameters of sub-groups, detected that these spaces are far from being isomorphic, with a size of possible local isomorphisms involving no more than 20 members. The remaining option, a social network which was not anonymized and could allow measuring distances was the co-author network. However, as explained before, its small density rule out a robust statistical analysis.

These results throw doubt on the use of social networks to describe knowledge recombination processes since they might fail capturing knowledge distances. People located at the periphery of the social structure might not have access to heterogeneous knowledge, or people at the socially dense core of an organization may be surrounded by highly heterogeneous knowledge, for example. The revision of this assumption of isomorphism needs not to be destructive but the opposite. The social dimension of innovations is undeniably crucial to explain collective process of knowledge creation. However, we must avoid falling in the simplistic approach of explaining knowledge heterogeneity by analyzing social networks. In the case of patentable technology or academic research, citation networks offer a valuable source of information about the knowledge content not only in patents and papers as their basic units, but in people and organizations as well. Summing up, a proper approach to study innovation as social driven should treat both dimensions as different and complementary in the explanation of such important phenomenon.

The interesting question to look at is what happens if the isomorphism is broken. Two symmetric configuration can be identified if this assumption is dropped. First, groups with short social distances but long knowledge distances among their members. This kind of configuration is the multidisciplinary or cross-functional group, a team composed by people with very different knowledge background for example. The diversity of knowledge involved is combined by the social proximity of their members. On the other extreme, isolated and socially distant groups working on the same knowledge area describe the symmetric scenario. This can be called duplicated components or redundancies if those groups belong to the same firm. A full analysis of these structures is developed in Chapter 3.

"Networks poised at the edge of chaos
can perform the most complex tasks"
Simon (1962)

Chapter 3

The Innovative Disorder

The Impact of Social-Knowledge Redundancies on Innovation

Abstract

This chapter challenges the idea of redundancies in organizational structures as duplicating costs that can only serve as a fail-safe. Instead, this chapter argues that they may positively impact on the organizational ability to innovate. Along with redundancies, cross-functional structures are also studied as their opposite configuration. By analyzing the dual structure of organizations, the knowledge base and collaborative network, redundant and cross-functional structures are identified. While the former structure is observed when people with similar knowledge are distantly located in the social structure, the latter happens when inventors with very different knowledge work together. This chapter argues that redundancies may play an important role when innovating with complex technologies which requires a systematic thinking. On the other side, cross-functional structures may enhance innovation when disruptive thinking is necessary. In order to differentiate these two cognitive abilities, I look at the interdependence of technologies. High interdependent technologies require trying and assessing many uncorrelated possibilities in order to find successful innovations; low interdependence requires being able to make radical changes of design in order to succeed. Using information on inventions patented by firms operating in the semi-conductor industry, empirical evidence supports the idea that cross-functional structures enhance the ability to innovate with low interdependent technologies, while organizational redundancies increase the capacity of innovating with complex ones.

3.1 Introduction

The presence of redundancies in a system is usually defined as the existence of more than one component performing the same function (Landau, 1969; Felsenthal, 1980). This duplication of functions provide reliability to the system by allowing partial failures without menacing the whole performance. Redundancy of components is a principle well known among engineers for designing fail-safe machinery as it is also omnipresent in living organisms as means of adaptation and self-reparation (Kauffman, 1993). Generally, either applied to engineering, linguistics, networks, information theory, or biology, redundancies imply having parallel mechanisms that perform the same task which, in turn, provides reliability to the whole system.

Organization theory is not an exception. Redundancies in organizational structures are also considered to provide reliability (Landau, 1969; Weick et al., 2008; Caldwell and Wang, 2010; Csaszar, 2013). Since firms decompose complex goals into simpler and articulated tasks performed by different sections, the existence of more than one section executing the same task prevents a general failure. However, redundancies can be easily stigmatized as they do not offer straightforward benefits though they do represent a duplication of costs (Staber and Sydow, 2002). They might be even explained as consequences of communication problems when different business units are unaware of similar efforts done by other business units, or as consequence of managerial incompetence by not clearly defining roles and competences (Felsenthal, 1980).

This chapter challenges this perspective arguing that redundancies might offer much more than a fail-safe organizational structure. Duplicated components are here argued to be a key characteristic of the organizational capacity to deal with knowledge. When tackling problems of great complexity, having independent organizational areas devoted to the same task might offer an advantage for finding a solution. Thus, instead of being a burden, redundant components might empower the organizational information-processing capacity.

In order to empirically assess the role of redundancies in organizations, the very concept of redundancy must be operationalized. Redundancies are defined as different and unconnected parts of an organization performing the same or similar function. This definition requires two dimensions to be detected: the dependency in the organizational structure *and* the similarity among functions. The first describes the connection level among organizational components. In this dimension, closely located components would be those that are strongly connected by a constant interaction and exchange of knowledge. On the contrary, sections or employees that rarely collaborate or consult each other are considered as distant located. The second dimension describes the similarity among tasks according to the similarity or complementarity of their goals. For example, the goals of designers and engineers in a car company are expected to complement more than the ones between designers and psicologists working for the human resources department. Considering these two dimensions, redundancies are observed when unrelated organizational parts pursuit similar goals. For example, if a group of engineers is devoted to solve the same problem of certain engine but in two independent teams without being aware of what the other part does, they are redundant.

This chapter uses a novel approach for capturing these two dimensions for analyzing the innovative activity of firms. On the one hand, distances in the organizational structure are measured in terms of inventors' location in the firm's internal collaborative network. This analysis does not consider the organizational structure as the formal design of the firm that establishes hierarchies, units or locations, but as the structure of collaboration among employees. On the other hand, since the function of inventors is to innovate in those technological areas they are expert on, function is identified with knowledge expertise. Therefore, the similarity among functions is measured in terms of expertise similarity. As it will be explained in the next sections, this similarity is captured as distances in a knowledge space where inventors are located according to what they know. Then, by combining the two dimensions, inventors who are distantly located in the collaborative structure but closely located in the knowledge space are considered to be redundant since they are researching on the same topics in parallel.

The operationalization of redundancies allows the definition of a mirror concept: cross-functionality. Those inventors that are closely located in the collaborative network though they are expert on very different technological areas represent a cross-functional structure. Since this chapter identifies function with expertise, this structure can be also called multidisciplinary. Contrary to redundancies, these structures are easily associated with innovation because of their capacity to bring together diverse knowledge.

Both redundant and cross-functional structures depict two different ways organizations process knowledge. As such, they are expected to differently affect the innovative capacity. None of them is assumed to be superior than the other one *per se*. Instead, they are expected to outperform the other in different scenarios. Thus, in order to distinguish their strengths, this chapter uses the interdependence of innovations as a key characteristic that reveal different cognitive abilities involved in innovating. Based on the adaptation of the NK model (Kauffman, 1993) to the innovation field by Fleming and Sorenson (2001), the interdependence is defined as the level of interaction among components in terms of contribution to the overall performance. High interdependent configurations make the design's performance very sensible to minor variations in any of its components. This increases the difficulty of finding the best possible configurations. Oppositely, low levels of interdependence make problems decomposable and easy to solve by parts. This helps finding the best configuration as well as it makes them very stable (Skellett et al., 2005).

Doing a different reading of the NK model, the level of interdependence of innovations and its success are here seen as revealing the cognitive abilities of the innovator. Different levels of interdependence demand different abilities to find better configurations. While searching for optimum designs with high interdependent technologies requires an exhaustive search, trying and evaluating many possible variations; improving configurations with low interdependence demands being able to radically change approaches. While the first ability is related to systematic thinking, the second is related to disruptive thinking. Since organizations do not randomly try possible configurations but instead they actively interpret technologies (Yayavaram and Ahuja, 2008), being successful in these scenarios reveals the cognitive ability of the firm.

Summing up, two organizational structures are defined (redundancy and cross-functionality) along with two different cognitive abilities required to innovate (systematic thinking and disrupting thinking). This chapter proposes that redundancies are expected to positively enhance the capacity for assessing many possible configurations (i.e. systematic analysis) when dealing with complex technologies. On the other hand, cross-functionality is expected to increase the ability to improve simpler technologies that require cognitive disruptions in order to find new configurations. Using patent data on firms operating in the semiconductor industry, these hypotheses are empirically tested with supporting results.

This research contributes to the innovation literature in three ways. First, it offers a simple operationalization of two major concepts in the management literature: cross-functionality and redundancy. While the first has long been studied, the second has not received the same attention, especially with regards to its role in innovation inside firms. The impact of cross-functionality is more intuitive than redundancies in this matter. Furthermore, both structures, redundant and cross-functional are here defined as two sides of the same phenomenon: how organizations articulate the knowledge (or function) space with the social structure. Second, this chapter also operationalizes two very distinct cognitive challenges an innovation may pose: those which require systematic processing and those which require disruptive thinking. Third, the operationalization of the previous concepts allows an *empirical* approach for studying the role of redundancies in firms, something that has never been done before. This way, this chapter challenges the undervaluation of the role of redundancies on organizations' performance.

In the following sections, the previous literature on these concepts is explored for proposing the hypotheses previously suggested. Afterwards, Section 3.2.2 discusses the nature of cognitive abilities for innovating. Section 3.3 describes the chosen data set and how interdependence is calculated. Later on, at Section 3.3.2 the operationalization of redundancy and cross-functionality is fully explained. However, details on the way metrics are calculated are disclosed at Section A.1 in the Appendix. Finally, after explaining the statistical analysis and results, Section 3.4 discusses their meaning and implications. Final thoughts are included in the conclusion.

3.2 Theory and hypotheses

An organization can be understood as a social arrangement that collectively processes information (Kogut and Zander, 1992; Radner, 1993; Nonaka, 1994; Bolton and Dewa-

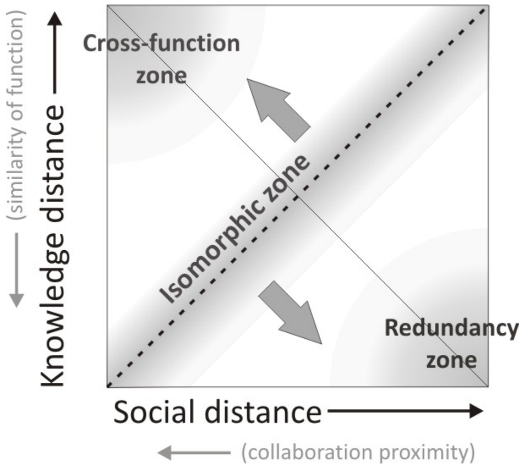


Figure 3.1: Conceptual representation of redundancy and cross-functionality

tripont, 1994; Kogut and Zander, 1996; DeCanio and Watkins, 1998). Complex problems are decomposed into simpler ones and then articulated into functional solutions throughout a social system that structures specialized sections (Grant, 1996; Palomeras and Melero, 2010; Nelson and Winter, 2009; Nonaka, 1994). Achieving this requires a social structure capable of transmitting complex knowledge and propelling collaboration among employees (Kogut and Zander, 1992, 1996). This way organizations amplify knowledge and processing capacity so they are able to deal with a level of complexity beyond individual cognitive limitations.

The information-processing structure of organizations combines two dimensions. On the one hand, it is a network of social interactions among employees. On the other hand, it is an articulated structure of different functions. These two dimension are expected to be "aligned" as organizations group workers according to their specialized function (Ethiraj and Levinthal, 2004; Grant, 1996; Hislop, 2013; Hendriks and Fruytier, 2014; Mintzberg, 1993). If workers assigned to the same task strongly collaborate with each other, while workers assigned to different tasks occasionally collaborate, it would be expected a correspondence between collaboration intensity and function combinability. In other words, similar functions would be associated with strong ties, while unrelated functions would be associated with weak or indirect ties. As long as these two dimensions are aligned, it does not make much sense to differentiate them. However, what would happen if there was not such a correspondence among them? Specifically, what are the consequences on innovation? Some researches have devoted attention to the informal social structure of organizations (Tortoriello and Krackhardt, 2010; Tortoriello, 2005; Tsai and Ghoshal, 1998; Tsai, 2000), others to the formal structure (Damanpour, 1996; Hansen and Løvås, 2004), some researches have focused on spatial location

(Hansen and Løvås, 2004; Singh, 2008), others on the knowledge structure (Yayavaram and Ahuja, 2008; Yayavaram and Chen, 2013). When it comes to innovation, the organizational capacity to combine knowledge plays a crucial role. However, previous research has focused either on one side or on the other, implicitly assuming that one reflects the other. This chapter argues that organizations combine knowledge through the interplay between the social and the knowledge structure, and this interplay does not imply a correspondence between these two dimensions.

When analyzing innovation, the function of an inventor can be identified with her/his expertise. This identification is expected to especially operate in those technological fields where the knowledge involved is advanced and not easily acquirable. Complex knowledge demands researchers to invest time and resources in order to acquire it, and therefore, they are expected to profit from their expertise by doing research on it. For instance, some industries, as the pharmaceutical or the semi-conductor industries, are known for developing collaborative ties with scientists and universities (Fleming and Sorenson, 2004) indicating the high level of education that inventors need for doing their job. Consequently, it is expected that only experts can research on advanced technological fields. Expertise and a researcher's function within the organization, then, converge as knowledge is more difficult to master. This identification of expertise with functions allows describing the similarity of functions as combinability of knowledge. Therefore, when studying innovation, "aligning" collaboration and functions within an organization would imply a correspondence between collaboration intensity and expertise combinability. After this specification, the previous question can be restated as what happens if experts in the same field do not collaborate or if experts in different fields do collaborate? What happens if similar knowledge is processed in parallel by different sections or if very distinct knowledge is combined by the same group?

The no-alignment of the collaboration network with the knowledge structure is expected to impact on the innovation capacity (Grant, 1996; Nonaka, 1994; Nelson and Winter, 2009). Since innovations can be understood as recombination of knowledge, it is natural to expect that the way knowledge is socially combined and distributed (collaborative network) and the propensity to be combined (knowledge structure) jointly affect the whole process. Therefore, the absence of a correspondence between collaboration proximity and knowledge combinability may describe different ways in which knowledge is collectively combined.

3.2.1 Redundant and cross-functional structures

In order to analyze the deviations from the correspondence between the collaboration network and the knowledge base, they will be described as metric spaces: the social space and the knowledge space. The collaboration network can be considered as a social space where distances among inventors are defined according to direct and indirect collaborative ties as well as their strength. In this space, those inventors who strongly collaborate would be considered to be closely located. Oppositely, those who do not know each other and do not share any common colleague but are indirectly connected would be considered to be distantly located (Wasserman and Faust, 1994). These distances would indicate the propensity of knowledge to be channeled among inventors (Owen-Smith and Powell, 2004a). The stronger the collaboration tie, the closer two inventors are in the social structure, the stronger their capacity to transmit knowledge. Thus, the whole set of distances among inventors would describe how knowledge is channeled and combined by the firm.

Inventors can be also located in a knowledge space according to what they know. As knowledge is combined, its distribution across the members of the organization can be also described in terms of distances. The complexity and scope of knowledge make organizations hire (Palomeras and Melero, 2010) or train specialists and then combine what they know throughout collective work (Carley, 1990; Ethiraj and Levinthal, 2004). This expertise stored in employees can be described in terms of combinability and then translated into a measure of distance. Types of knowledge that are usually combined can be considered as closely located in a knowledge space. Those that are never combined are more distantly located though their distance is defined according to indirect connections as it was a network (Yayavaram and Ahuja, 2008; Yayavaram and Chen, 2013).

These two spaces generated by an organization, the social and the knowledge space, are expected to be isomorphic (Chapter 2). Distances among the same elements within these two spaces should be a monotonic transformation: we would expected that combinable knowledge should be combined by close collaborators, while less combinable knowledge should be combined by inventors who are more distantly located in the social space. Differently stated, the ismorphism between these two spaces implies that those employees assigned to similar functions should closely collaborate while workers with very different functions should not collaborate on daily basis but eventually. When this positive relation among distances does not hold, two general structures can be identified as deviations.

One scenario is the case of redundant structures. When experts in similar knowledge do not collaborate with each other but are distantly located in the collaborative structure, the organization is dealing with the same knowledge in parallel. Using the scheme portrayed in Figure 3.1, redundancy is defined as the deviation from the isomorphism between the social and knowledge space where distances in the first are larger than distances among the same elements in the second space.

This operazionalization of the concept of redundancy captures the level of social disconnection and the level of overlapping of knowledge. The larger the deviation, the stronger the redundancy. Since inventors belong to the same organization, knowledge is expected to eventually flow among most of the members (Cowan and Jonard, 2004, 2009). Thus, the social distance among them measures the time or effort it takes to flow (Cowan et al., 2004; Singh, 2005). This continuum allows differentiating the level of independence of different sections. Longer social distances exaggerate the redundancy. Proximity of knowledge also allows measuring the intensity of the redundancy. Independently performed functions can be identical, similar, or potentially combinable to some degree.

As far as I am concerned, the innovation literature has not dealt directly with the concept of organizational redundancy. It has been theoretically discussed (Csaszar, 2013; Landau, 1969; Nonaka, 1994; Weick et al., 2008), but it has never been operationalized, empirically captured and related to innovation. Although, some related concepts have been studied in the innovation literature.

A long studied idea on the literature that might be related to redundancies is *knowledge overlapping* among inventors. This concept has been recognized to allow and improve communication and therefore it can affect the combination process (Ethiraj and Levinthal, 2004; Mowery et al., 1998; Nonaka, 1994; Thayer and Barnett, 1996; Uzzi and Spiro, 2005). The approach used in this chapter embraces this idea but goes beyond. First, knowledge distances among inventors include those inventors whose knowledge does not overlap. Different knowledge sets may not overlap but still be combinable to different degrees among them. Secondly, the literature that studies the effects of overlapping knowledge only considers the case of interacting inventors. In other words, it only analyzes the case of socially close inventors with different degrees of knowledge overlapping. This chapter allows the social distance to vary beyond direct interaction. For instance, the expertise of two inventors may totally overlap but they independently operate within the organization so it does not have any direct effect on them.

Another concept in the literature that can be considered as similar to redundancies is *organizational slack*. The comparison may come from the fact that both of them represent a departure from efficiency. Slack is defined as "the pool of resources in an organization that is in excess of the minimum necessary to produce a given level of organizational output" (Nohria and Gulati, 1996). Actually, the use of the word "redundancy" sometimes is associated with this concept of mere excess. This chapter takes a more strict definition by considering functionality. Within a system, redundant components are those that are not directly related but assigned to the same function. This definition would be a particular case of slack since it can be considered as an excess of resources (with the same purpose. However, slack is a broader definition that considers other kind of resources as capital, number of employees, capacity, etc. The distinction between these two definitions is relevant to this research since organizations are here analyzed as mechanisms that combine knowledge where redundancies depict a particular configuration rather than a vague excess of resources. In fact, redundant components might not represent an excess at all if seen as a necessary condition for some processes.

Thirdly, the study of geographical location and innovation can also be related to redundancies. Not only social distances may mimic geographical distances (Feld, 1981), but also knowledge flows have been found to depend on geographical proximity (Fleming and Waguespack, 2007; Fleming and Frenken, 2007). Therefore, when organizations are spatially dispersed, the study of how knowledge is combined throughout interaction of different units (for instance Zaidman and Brock (2009) and Singh (2008)) could be related to this chapter's topic. However, this chapter directly focuses on collaborative ties which are not necessarily related to geographic position. Furthermore, although the geographical approach could be used (only for some firms) in order to study the effects of redundancy on innovation, it has never been done in the literature.

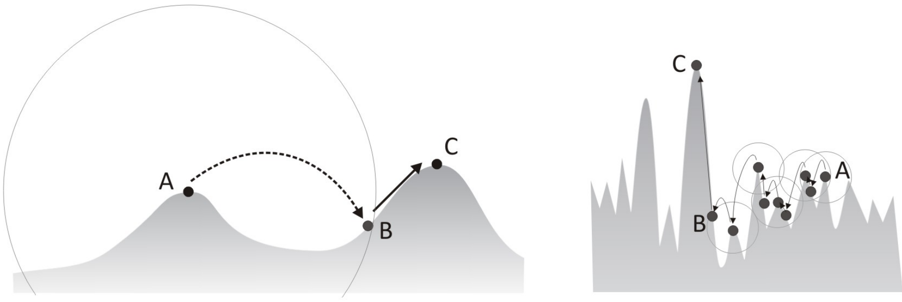
In contrast with some of the aforementioned approaches on the innovation literature, redundancies are here considered as a characteristic of mechanisms with independent components performing the same function. Living beings as well as their DNA are known to have redundant components in this sense (Kauffman, 1993). Complex machinery as airplanes are also designed fail-safe with components that perform the same function in order to prevent catastrophic failures. The general principle is that redundancy provides reliability to modularized systems. This, in turn, allows living organism to adapt and evolve as well as it makes machinery safer. Those scholars in management who have analyzed redundancies in this fashion, have proposed that redundancies might have the same purpose in organizations: reliability (Caldwell and Wang, 2010; Csaszar, 2013; Landau, 1969; Staber and Sydow, 2002; Weick et al., 2008). This chapter proposes that organizational redundancies might represent more than reliability. They might describe the configuration of a processing machinery that uses the parallel activity to tackle problems of certain complexity.

The operationalization of redundancies adopted here allows the identification of cross-functionality as its mirror concept. If the collaborative network describes the social space, and the knowledge structure describes the knowledge space, cross-functionality would be defined as the deviation from their isomorphism but in the opposite direction of redundancy. As Figure 3.1 shows, cross-functionality would indicate a large heterogeneity of knowledge within groups with close collaboration ties. Since this chapter focuses on innovation, expertise and function are identified and therefore, cross-functionality is here identified with multidisciplinary structures. This identification does not necessarily hold if we consider the entire organization since experts in the same area could be assigned to different functions. However, in high technological industries, where knowledge is complex and difficult to acquire, inventors are expected to research on their area of expertise (Kim et al., 2004).

The relation between multidisciplinary groups and its innovative capacity is intuitive mainly because innovations are considered as combinations of knowledge. Thus, increasing the diversity of knowledge embedded in people would likely increase the scope for combining. Unlike redundancies, cross-functionality has a direct relation with innovation and therefore the literature has devoted more attention to it. In general terms, bringing together diverse knowledge through collaboration is an idea so essential to the innovation literature that it is pointless to mention specific research on this topic.

Although this is a largely accepted idea (Ford and Randolph, 1992), this chapter differs from previous research by considering cross-functionality as a possible configuration of the organizational cognitive machinery. Furthermore, similar to the operationalization of redundancies, cross-functionality is not considered here as a dichotomous characteristic of a certain structure. Cross-functionality is measured in a continuum indicating that some structures are "more" cross-functional than others. The more diverse the knowledge involved and/or the closer the collaboration of people, cross-functionality is expected to increase its impact on innovating.

Figure 3.1 represents the two deviations from the isomorphism. While one axes depicts social distance, the other one indicates knowledge distance. The 45° line draws the isomorphic zone where the social distance and the knowledge equally increase. Redundancies could be observed when social distances are larger than knowledge distances. In Figure 3.1 this corresponds to the bottom-right of the plot. The other deviation from



Low interdependence of components High interdependence of components

Figure 3.2: Simple representation of different roughed landscapes according to the interdependence of components (K) in the NK model (assuming N constant).

the isomorphism happens on the top-left corner, when different knowledge is socially mashed up by a cross-functional structure.

The two deviations from the isomorphism, redundancy and cross-functionality, may affect the capacity for innovating in different ways. While redundant structures process the same knowledge in parallel, cross-functional structures process different knowledge in the same group. The impact of such distinct configurations for recombining knowledge is expected to be non-trivial. There are no reasons to think that one of them *necessarily* prevails over the other. However, due to its deeply distinct nature, both structures are expected to impact differently on different kind of innovations.

In order to analyze this possibility, the following section introduces a novel way of understanding the cognitive processes behind innovating. While the former paragraphs have discussed two distinct organizational configurations that emerge from braking the strict correspondence between collaboration and knowledge proximity, the next section radically switches the topic towards the cognitive abilities require for innovating according the level of interdependence of its components. This jump in the argumentation is necessary for merge the analysis of organizational configurations and the required cognitive abilities for innovating for explaining the the success of inventions.

3.2.2 Interdependence as a cognitive challenge

Technologies can be described as composed by several interacting parts or assemblies (Arthur, 2009). For instance, while a microprocessor assembles a central processor unit (CPU) into an integrated circuit on a semi-conductor plate, an integrated circuit is formed by assembling millions of transistors and a semi-conductor material can be formed by combining silicon and carbon (silicon carbide). The different parts composing a technological device perform different functions that contribute to fulfill the overall purpose of such device. However, these parts are not independent by they interact with each other affecting their functions. This interaction is what Kauffman (1993) called *interdependence* and it plays a crucial role in determining the nature of the problem of finding successful configurations.

The role of interdependence in determining the difficulty of finding successful designs was modeled in the *NK*model (Kauffman, 1993). Originally proposed for evolutionary processes by considering the number and the interdependence of genes, this model explains the fitness of genomes as a rugged landscape. Different positions in this landscape represent different genomes, where distances among them describe the similarity of those configurations, and where altitude represents different levels of fitness. The level of interdependence among genes, then, determined the ruggedness of the fitness landscape. In other words, the level of interdependence determines how much the performance of configuration varies with minor alterations of its components. High levels of interdependence imply that minor alterations on the configuration lead to abrupt and uncorrelated changes on the performance. We can think of this case as a Rubik's Cube (the famous 3D puzzle) where changing one color of one piece implies affecting the color of several others. Low levels of interdependence, on the contrary, describe configurations where minor changes lead to minor alterations on the performance and with a clear correlation. A normal puzzle has zero interdependence given that individual pieces do not affect the contribution of the rest.

The logic of the *NK*model was adapted to the technology field by Fleming and Sorenson (2001) finding supporting evidence on patent data. While their goal was to inquire whether this model adapts to technological innovation, this chapter proposes another reading of this model based on their adaptation. With regards to interdependence, the *NK*model basically predicts two relations. First, the number of local optima increases with a decrease in their average altitude. This implies smaller basins of attraction. Second, the absolute maximum is expected to increase as interdependence does. This statement, combined with the previous one, indicates a less correlated function and an increase in the variance of the altitude across the entire landscape. When translated to technology, interdependence would reveal not only the nature of the landscape but also the nature of the person navigating that landscape.

The NK model (Kauffman, 1993) assumes evolution to be a blind searching process over a fitness landscape that seeks for high locations by random mutations (Back et al., 1997; Holland, 1975). Thus, by determining the nature of the landscape, interdependence also determines the expected success of a adaptative walker as it blindly navigates it. However, if we drop the assumption of blind navigator, interdependence only determines the nature of landscape but not the expected success. Innovation differs from evolution in the sense that is not a random process but innovators actively interpret the landscape (Fleming and Sorenson, 2004; Yayavaram and Ahuja, 2008; Yayavaram and Chen, 2013). Therefore, by considering an active navigator, the success she or he achieves is independent of the interdependence it faces. Then, the success of an innovation together with its level of interdependence reveal the cognitive abilities of the organization.

This chapter assumes that inventors combine components by using their cognitive criterion to assess the better configurations ex ante. Using the metaphor of the technological landscape where the level of interdependence determines its ruggedness, two cognitive abilities can be identified. While high success with low interdependent technologies reveal innovative groups with a great capacity for disruptive thinking, high success with high interdependent technologies reveal groups with a great capacity for methodic or systematic processing.

Low interdependence among components describes a smooth landscape as depicted at the left panel of Figure 3.2. This landscape is characterized by large basins of attraction, few and distant local optima, and a highly correlated space. Optimal configurations are easy to find, although they are very stable and therefore they put up more resistant to change from one optimum to another. As Figure 3.2 shows, once located in an optimum, the innovator should jump a long distance in order to escape the hill that leads to its location. Radically changing the design of a very stable configuration involves being able to abandon conventions, going beyond familiar ground, making cognitive jumps, thinking out of the box. Innovating with low interdependent technologies is here assumed to demand disruptive thinking (Csikszentmihalyi, 1997) in the sense that the mind behind the invention must disrupt the current approach.

On the other extreme, high interdependent technologies describe a rugged landscape as the right panel of Figure 3.2 shows. This landscape is characterized by small basins of attraction, many and close local optima, and an uncorrelated space. This indicates how much the performance of an invention can vary with minor alterations of its design. The difficulty of improving highly interdependent configurations is of a different nature compared to the previous case. The absence of spatial correlation among changes and performance makes finding the best configurations more difficult.

A shortsighted navigator would be easily trapped in a local optima. As Figure 3.2 portraits, finding the best design does not require a great ability to jump but it does require trying and evaluating many locations. Then, innovating with high interdependent technologies is here assumed to demand a great capacity of trying and evaluating many possibilities, an ability that this chapter calls *systematic thinking*. Solving the Rubick's cube does not require disruptive thinking but the capacity of trying and keeping track of hundreds of moves before loosing hope.

Summing up, the level of interdependence of components captures two abilities necessary for innovating. An inventor dealing with low interdependent technologies must think disruptively in order to jump to another basin of attraction. Once there, marginal improvements easily lead to the new local optimum. In the rest of the chapter, this ability will be called *disruptive thinking*. On the contrary, high interdependent technologies do not require long jumps but they do demand the capacity to canvass large set of possibilities in order to find a more successful configuration. This ability will be called *systematic thinking* since it involves dealing with a lot of attempts. Therefore, while disruptive thinking applies whenever solutions are easy to find but difficult to change, systematic thinking applies when solutions are easy to change but difficult to find.

3.2.3 Redundancies and systematic thinking

According the NK model, high interdependent technologies present three challenges for finding the best configurations: the absence of clues (one possible configuration does not lead to better ones), a low expected value of success, and the number of local optima. However, they also provide fewer but much more successful configurations than low interdependent technologies. The problem is to find them.

Organizational redundancies, as they were defined in Section 3.2.1, may improve the firm's performance when dealing with high interdependent technologies. First of all, the spatial overlapping of different units when exploring the landscape does not affect the performance due to the lack of spatial correlation. Searching in parallel on close locations may lead to two different extremely successful designs.

Secondly, organizations with redundancies may be capable of avoiding sub-optimal solutions. The analysis of complex adaptative systems (CAS) proposes that the presence of redundancies might generate some "noise" that helps the system abandoning stable configurations and, therefore, evolving (Sherif and Xing, 2006; Von Foerster, 1984). The different perspectives or approaches adopted by redundant units may eventually affect each other and generate perturbations that can push inventors to leave their local optima (Marengo et al., 2000). As Ethiraj and Levinthal (2004) stated, ”moderate amounts of such self-perturbation have a useful property of encouraging search and preventing premature lock-in to inferior designs”. In rugged landscapes, the possibility of being trapped in local optima is large but it needs less effort to abandon this position than in smooth landscapes. Therefore, loosely connected units that explore the same problem may not only increase the chances of finding good solutions, but they also may prevent other units from being trapped in sub-optimal ones.

Furthermore, assuming a convergence of what people know proportional to the intensity of their collaboration ((Acemoglu et al., 2010; Cowan and Jonard, 2004; Holme and Newman, 2006; Owen-Smith and Powell, 2004b)), the presence of redundancies may prevent the firm from a search bias in the exploration. Since patterns disappear as interdependence increases, the landscape must be swept without ignoring the possibility of high optima in unexpected places. In the words of Sherif and Xing (2006), “the process of parallel exploration makes it easier to incorporate new information”. Therefore, redundancies may help avoiding the prevalence of cognitive traps or paradigms that can rule out large areas from the technological landscape without evaluation.

In sum, high interdependent technologies do not allow tracing routes but they force canvassing entire areas for finding the best solutions. This requires avoiding biases in the exploration and being trapped in local optima, and a great capacity for carrying repeated trials. Organizational redundancies may help this process by increasing the randomness of searching, generating noise across loosely connected units, and preventing exploration from biased approaches by generating alternative approaches with different solutions. Because of these reasons, the following hypothesis is stated.

Hypothesis 1. *The presence of redundancies in an organization positively affects its ability for successfully innovating with high interdependent technologies.*

3.2.4 Cross-functionality and disruptive capacity

When dealing with low interdependent technologies, the capacity of finding successful configurations lies in the ability to break well stated designs and jump to totally different ones. A low level of interaction among components makes problems easier to solve since minor variations in the configuration point into the direction of better configurations. However, they also pose more resistant for innovation. The stability of successful configurations demands disruptive thinking in order to find better configurations.

If we assume inventors as having a limited cognitive capacity to analyze long ranges of variations of a design, the only way of covering long areas is by teams consisting of people with disparate knowledge. Then, since organizations use the social space for combining knowledge, it would be expected from cross-functional or multidisciplinary groups to do better in going far from familiar ground (Mumford and Gustafson, 1988; Taylor and Greve, 2006; Quinn, 1985).

The relation between multidisciplinary groups and disruptive thinking does not need much argumentation. One of the most known definitions of innovation describes them as combinations of knowledge (Schumpeter, 1939). The very definition implies diversity of combinable elements. Therefore, knowledge heterogeneity is generally accepted in the innovation literature as necessary for innovating (Ethiraj and Levinthal, 2004; Leonard and Straus, 1997; Love and Roper, 2009; Perry-Smith, 2006; Reagans and Zuckerman, 2001; Simon, 1985; Sytch and Tatarynowicz, 2013). In this case, the role of the organization is to foster the communication and collaboration among people with different knowledge backgrounds and expertise. The clash of their ideas, approaches and knowledge of different areas in the technological landscape may create enough conflict to break ideas and abandon the stability of a local optima in the smooth landscape.

Hypothesis 2. *The presence of cross-functional structures in an organization positively affects its ability to successfully innovate with low interdependent technologies.*

3.2.5 Size and structure

Although there is ample research analyzing the relation between organizational size, innovation and complexity of structure[1], the following two hypotheses are based on mere logic. The correspondence between social and knowledge distances cannot hold independently from the firm's size. Maintaining this correspondence and expanding the firm in terms of members would imply a proportional expansion on the knowledge space. This may be as undesirable as impossible. Firstly, expanding the knowledge base too much might surpass the firm's goals. Secondly, embracing broader areas of technology can eventually overwhelm the processing capacity of the firm. Therefore, if organizations increase their processing capacity by hiring more inventors, their areas of expertise are expected to overlap. On the contrary, when firms are small, multidisciplinary structures might expand the knowledge base and increasing the capacity of

[1] Damanpour (1996) offers an exhaustive review and analysis on these matters.

Class	Relative %	Cumulative %	Name
438	16.10	16.10	SEMICONDUCTOR DEVICE MANUFACTURING: PROCESS
257	9.58	25.68	ACTIVE SOLID-STATE DEVICES
365	6.38	32.06	STATIC INFORMATION STORAGE AND RETRIEVAL
327	4.26	36.32	MISCELLANEOUS ACTIVE ELECTRICAL NONLINEAR DEVICES, CIRCUITS, AND SYSTEMS
439	3.18	39.50	ELECTRICAL CONNECTORS
326	2.67	42.17	ELECTRONIC DIGITAL LOGIC CIRCUITRY
711	2.46	44.63	ELECTRICAL COMPUTERS AND DIGITAL PROCESSING SYSTEMS: MEMORY
375	2.24	46.87	PULSE OR DIGITAL COMMUNICATIONS
710	2.24	49.11	ELECTRICAL COMPUTERS AND DIGITAL DATA PROCESSING SYSTEMS: INPUT/OUTPUT

Figure 3.3: Most populated main technological classes in the semiconductor industry

generating successful innovations. Because of these reasons, the following hypothesis are proposed:

Hypothesis 3. *The impact of redundancies on a firm's ability to innovate increases with the firm's size.*

Hypothesis 4. *The impact of cross-functional structures on a firm's ability to innovate increases as the firm's size decreases.*

The following sections explain in detail the chosen theoretical framework as well as the variables proposed to capture the presence of redundancies and cross-functional structures in organizations. The empirical setting is also described in order to propose a statistical test for both hypotheses. Results are discussed in their scope and meaning afterwards. Finally, the Appendix includes a detailed explanation on the proposed measures and an alternative statistical analysis by using a different constructed variable for capturing redundancies.

3.3 Empirical analysis

Testing the previous theoretical ideas requires a proper data set. Specifically, it is needed information about inventions, their components, their level of interdependence and their success. Information about the organizations which have generated these inventions is also necessary. The analysis requires building the social and the knowledge space that organizations generate in order to study how they affect each other. All this data can be obtained by the analysis of patented inventions.

Patents are broadly used in the innovation literature mainly because they provide accurate and detailed information on an invention's design, inventors, owner firms, dates, previous inventions they are based on, other inventions that are based on it, and type of technologies used, among others. Since patenting an invention is a costly process and it is mainly used for legal means of protecting intellectual property, not all inventions are registered. This is one of the reasons of this chapter to focus on the semi-conductor industry since firms operating in this sector extensively appeal to patents for protecting their inventions. Furthermore, these firms not only patents most of their inventions but they also produce many patents as they focus efforts on innovating. The high research and development intensity (R&D) on this sector reveals the large dependency on its capacity to innovate in order to succeed (Yayavaram and Ahuja, 2008).

Another useful characteristic of the semi-conductor industry is the complexity of the knowledge involved in their technology. Electronics is a technology branch that mixes several scientific fields. This makes organizations a valuable unit of analysis given that this industry necessarily requires the collaborative work of specialist in multiple fields. Probably, organizations are the only locus where innovation may take place in this sector, and how they structure knowledge and collaboration would have a significant impact on their performance.

Those firms operating in the semi-conductor industry were identified on the Compustat Database (firms traded in the U.S. stock market) among active and non-longer existent companies with the SIC industry code 367, Electronic Components and Accessories. This code includes Printed Circuit Boards (3672); Semiconductors and Related Devices (3674); Electronic Coils, Transformers and Other Inductors (3677); Electronic Connectors (3678); Electronic Components (3679); and Electronic Components and Accessories (3670). Regarding the innovative activity, their patents were identified by consulting the NBER[2] U.S. Patent Citations Data File. Once firms and patents were matched, complementary information was obtained both from the Harvard Dataverse Network's *Patent Network Database* and the official website of the United States Patent and Trademark Office (USPTO) for each patent in the database. Several algorithms checked, corrected and completed the *Patent Network Database* by the systematic and direct consultation of the public on-line data base.

Following the approach of Yayavaram and Ahuja (2008), among all patents produced by semi-conductor firms, they were only selected those belonging to the 100 most populated technological classes within the data set (out of a total of 337 classes). This is done in order to narrow down the amount of technological classes involved in the analysis. These 100 technological classes represent 96.3% of the sample. As Table 4.2

[2]The National Bureau of Economic Research

shows, there is a high concentration on those technological classes directly related to the semi-conductor industry: only 9 technological classes accumulate 49.11% of the selected patents. The rest of classes are considered in order to include more diversity in types of knowledge components.

The USPTO is the ultimate source of information for this empirical analysis. Its classification system divides technological classes into sub-classes which are also hierarchically structured into 9 levels. In order to expand and capture enough diversity of technological characteristics, 3 levels of classification are considered: classes, level-1 sub-classes ("aggregated sub-classes" from now on), and sub-classes at maximum detailed level (simply "sub-classes" in the rest of the chapter). The final database contains more than 100,000 patents, 149 firms in 27 years, information from 1976 to 2010, 100 technological major classes, 1,836 aggregated sub-classes and 22,344 sub-classes.

3.3.1 Interdependence

As explained before, the analysis of the kind of intellectual challenges innovations may pose is based on the adaptation of Kauffman's NK model to the technological landscape by Fleming and Sorenson (2001). This approach considers patented inventions as the combination of technological-components represented by the technological sub-classes they are classified in. The USPTO's long list of hierarchically structured classes is intended to differentiate different technologies. Then, whenever a patent is classified in more than one sub-class, this invention is interpreted as it is combining the knowledge represented by those categories.

The interdependence among an invention's components (the parameter K of the Kauffman's model) is defined based on the characteristics of its components. Following Fleming and Sorenson (2001), before measuring the latter, the *ease of combination* of technological classes must be defined and measured. Considering the entire universe of USPTO's patent data set, for certain technological sub-class it is counted the number of sub-classes it has been historically assigned with, and the number of patents classified on it. The ratio between them indicates the average number of sub-classes the focal one has been assigned with. This number indicates how easy to combine is the focal sub-class. The larger it is, the less resistant the focal sub-class poses to be combined. In other words, technological sub-classes that are usually assigned to patents along with many others sub-classes are assumed to be easily combined. Then, the *ease of combination* of class c on year y is calculated as the mean of the number of classes class c has being classified with, in all patents that contain it with application date before year y. Equation 4.1 describes it:

$$E_{c,y} = \frac{\#\text{classes combined with class } c}{\#\text{patents classified in class } c} \tag{3.1}$$

Given the set of components that conforms the invention, Kauffman (1993) defines interdependence of components as how many components are affected in their contribution to the overall performance by the variation in one of them. This metric, although imperfect, captures this idea: those sub-classes that are usually combined with many others are assumed to exhibit low interdependence since they have shown low resistant to be combined.

For each 22,344 sub-classes in the data set from 1980 to 2003, their *ease of combination* was calculated by using the whole data set of patents contained in the Harvard Dataverse Network's *Patent Network Database* discarding design (D), statutory invention registration (H), plant (P) or reissue patents (R). The data set used to calculate this values is composed by a total of 3,984,771 patents classified in 17,445,405 classes/sub-classes. Since technology changes over time, the ease of combination is calculated yearly by considering all patents with application date earlier or contemporary. It is the application date, instead of the granting date, the one considered to date the patent since it better indicates the moment the inventions was finished and the combination of knowledge done. The granting date, on the other hand, informs the moment a patent is approved by the USPTO and it can be much different from the previous one. Considering application dates does not mean considering applications. All patents considered in this chapter are all granted patents.

Once the ease of combination is calculated for all sub-classes along all years, the interdependence of patents is measured according to the sub-classes involved. Considering a patented invention p as a combination of N sub-classes, then, the K parameter is calculated as the inverted average of the ease of combination of those sub-classes (Equation 4.2).

$$K_p = \left(\frac{\sum_{c=1}^{N_p} E_{c,y}}{N_p} \right)^{-1} \tag{3.2}$$

Those patents classified in technological sub-classes that have shown to be hardly combinable will score high in K. This estimation of a patent's interdependence was proposed by Fleming and Sorenson (2001) and replicated by other researchers. As they empirically proved, this parameter captures the expected behavior of the NK model applied to inventions. Both high levels and low levels of K showed lower expected citation rates on patents. Furthermore, high level of K showed a higher variance of performance as indicated by the model. These empirical finding hold in the data set used in this chapter when strictly replicating their approach.

3.3.2 Redundant and cross-functional structures

When organizations recruit inventors, they create social relations that enhance the transmission of knowledge and therefore the possibility of combining their expertise in innovations. By using information registered in patented inventions, it is possible to observe the social and knowledge space inventors are located in. Then, the comparison of these two structures will reveal the presence of redundancies. Specifically, they are observed whenever socially distant inventors research on similar technological areas. The opposite scenario, cross-functionality, is observed when socially close inventors have very different areas of expertise. This Section explains how both the social and the knowledge spaces are constructed, as well as how the presence of redundancies and cross-functional structures is measured. For the sake of simplicity, a much more detail description of this methodology is presented in section A.1 at the Appendix.

The social space

The web of interactions among employees created and fostered by an organization generate a *social space*. All the network's components and interactions jointly determine the capacity to transmit and collectively process information (Kogut and Zander, 1992; Radner, 1993; Bolton and Dewatripont, 1994; Kogut and Zander, 1996; DeCanio and Watkins, 1998). While the intensity of social relations indicates the amount and complexity of knowledge it can be transmitted between the involved persons, the structure and characteristics of indirect relations reinforces or weakens this capacity.

The social space describes the resistance for knowledge to be transmitted across people and therefore, it directly affects the functioning of firms. The design of an organizational chart is an attempt to control the flow of information across employees. However, the complexity of the social structure largely exceeds the formal structure (Tsai, 2001; Guimera et al., 2006b). Because there is always collaboration, advice, exchange of ideas or support outside the formal channels, the whole underlying network of interactions should be considered when analyzing a firm's innovative capacity.

Nonetheless, capturing all types of social interactions among an organization's inventors is nearly impossible because of the difficulty of defining and observing relevant social interactions. This chapter follows a broadly used approach, the co-authorship of inventions as officially registered by the USPTO (Allen, 1984b; Guler and Nerkar, 2012; Singh, 2005). Co-authorship describes close collaboration among inventors. When developing an invention, authors are assumed to have a strong tie capable of heavily transmitting knowledge. Co-authorship ties also allow the differentiation of intensities by considering the number of times two inventors have co-authored a patent. Thus, inventors and their co-authorship ties can be combined into a co-authorship network that approximately describes the social space where knowledge flows across.

This approach is not free from deficiencies. In the first place, it might not capture all strong ties among inventors. Even though co-authorship ties clearly indicate strong interactions, not all strong interactions are necessarily captured by co-authorship. In the second place, this approach rules out weak ties, those of occasional nature, unable to transmit tacit, uncoded or high complex knowledge, but capable of bridging parts of the social map and transmitting information that can play a major role when innovating (Granovetter, 1973).

However, although these limitations, this approach is still relevant for the purpose of this chapter. First, the wide use of co-authorship ties in the innovation literature shows the valuable information it captures on the social structure of inventors. Second and more importantly, for the sake of this chapter's analysis this information should provide enough detail. The complexity of knowledge involved in the semi-conductor industry requires strong channels of communication in order to transmit and collaborate. Since this analysis seeks to detect constant and general structures of redundancies and not to focus on those occasional or incapable of generating a deep flow of knowledge, the structure of co-authorship ties on the semi-conductor industry should suffice to capture the social space organizations use to empower research.

Directly downloaded from the USPTO on-line data base, the names of inventors of each patent are considered to construct the co-authorship network. Mapping this network usually presents the problem of matching names, a variable that can easily misidentify different elements. By considering the internal social networks of firms during 3-year windows, this possibility is significantly reduced. Using two different metric distances in order to increase the accuracy of identification, the full list of inventors and co-authorship were obtained for each firm for every 3-year window period from the information provided by patents. Then, in order to capture the presence and intensity of collaboration, the adjacency matrix of the network is built by counting how many patents are co-authored by every possible couple of inventors (Equation A.2 in Section A.1.1 at the Appendix).

The stronger the collaboration intensity, the closer co-authors are considered to be in the social space. The maximum intensity registered in the adjacency matrix is established as the closest direct distance between inventors. Whenever inventors have co-authored at least once, a direct distance is calculated as the maximum intensity of the adjacency matrix minus the intensity between them. In order to set the minimum direct distance as 1, all direct distances are added 1 (Equation A.3 in Section A.1.1). For those inventors that have not co-authored any patent, there is not a direct distance defined. Instead, indirect distances are calculated. Using the information on direct distances defined by co-authorship ties, the matrix of distances is calculated among all inventors as the shortest path on the collaborative network (Equation A.4 in Section A.1.1). Finally, a symmetric, definitive positive and integer square matrix describes collaborative distances among all inventors in the social space (Equation A.5 in Section A.1.1).

The knowledge space

If the social space captures how information and knowledge flows across people, the *knowledge space* describes how similar and combinable is what people know. While a social interaction opens the channel for exchanging knowledge, the dissimilarity among what the participants know determines how easy to transmit and how combinable it is. This dual structure of firms is crucial for innovation. While individuals specialize in small knowledge areas, organizations embrace much larger areas by collapsing distances on the social space.

It is possible to observe what inventors know according to the patents they have authored. Since these are classified by the USPTO according to the types of technologies involved, patents reveal what kind of knowledge the invention's authors master. This information allows measuring the similarity among areas of expertise of inventors and therefore they can be located in a knowledge space.

The approach used to build and locate inventors in the knowledge space differs from the one used for the social space. Following Yayavaram and Ahuja (2008), first I build the knowledge base for each firm each year. For doing this, all patents produced by the firm during the 3-year windows before the focal year are selected. By considering the technological aggregated sub-classes where patents are classified in, I build the coupling network. The knowledge base, then, is the network of technological sub-classes that certain firm has used during a 3-year period. The ties of this network are defined as the number of times certain couple of technological sub-classes have been coupled by a patent. Differently explained, if a patent is classified in technological classes x and y, it indicates that this invention is combining or *coupling* those technologies represented by classes x and y. The more patents couple those technologies, the stronger is the belief of the firm on that combination as a suitable and successful one (Yayavaram and Ahuja, 2008).

Following the same steps than with the collaborative network, first the adjacency matrix of the knowledge network is constructed (Equation A.8 at Section A.1.2). Differently to the social network, in this case each position does not correspond to inventors but to technological sub-classes. This matrix describes the coupling intensity among them. Based on this information, direct distances among sub-classes are calculated the same way they were calculated in the social network. Then, the whole matrix of distances is obtained by finding the shortest path across the network (Equation A.9 at Section A.1.2).

After obtaining the distance matrix among sub-classes, a multidimensional scaling is performed in order to place all technological aggregated sub-classes in a M-dimensional euclidean space such that distances among them are preserved (Equation A.10 at Section A.1.2). The set of coordinates each technological sub-class has in this euclidean space reproduces all distances with the rest of technological aggregated sub-classes. Once generated, inventors are located in the same space according to the patents they have authored. Specifically, an inventor is located in the weighted average (Equation A.14) of the coordinates of the technological sub-classes they have being classified throughout her/his patents.

Finally, once inventors are located in the knowledge space, distances among them can be easily obtained. A square, symmetric and definitive positive matrix describes all knowledge distances. Differently to social distances, in this case distances indicate the dissimilarity of areas of expertise.

Combined structure

After obtaining the full set of distances in both the social and the knowledge space among a firm's inventors during a 3-year window, it is necessary to define a way of capturing the presence of redundancies and cross-functional structures. Figure 3.4 helps with a simple example. At Panel I and II both spaces are represented with a group of

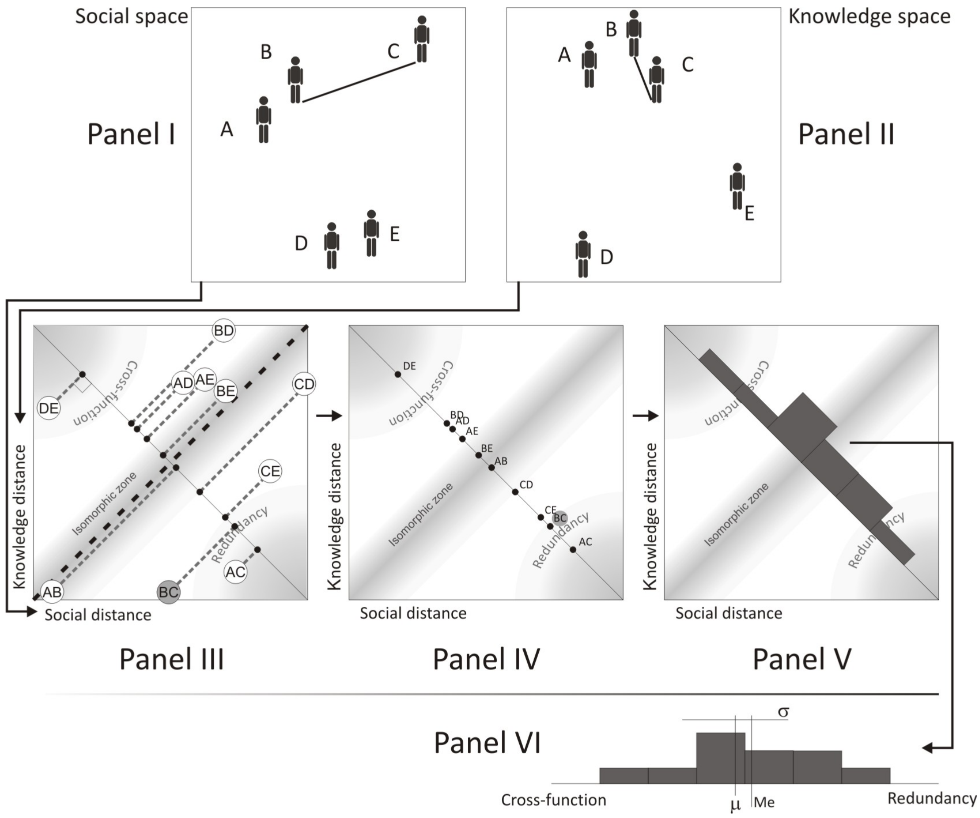


Figure 3.4: Panel I: Inventors located at the social space. Panel II: The same inventors located at the knowledge space. Panel III: Scatter of social distances versus knowledge distances and the orthogonal projection. Panel IV: Projections of all possible distances and their distribution. Panel IV: Histogram of projections in the orthogonal space to the isomorphic line. Panel VI: Histogram of projection and position and dispersion measures.

5 inventors named with letters from *A* to *E*. Every possible couple of inventors (a total of 5*(5-1)/2=10 couples) is associated with two distances: one in the social space and one in the knowledge space. Inventors *B* and *C* at Figure 3.4, for example, are far in the social space but close at the knowledge space. This pair of distances is plotted at Panel III with the label *BC*. This last panel scatters all possible pairs in their double distance on both spaces. As it can be seen, *BC* measures low on the knowledge distance axes and longer on the social distance axes. In other words, they do not each other but they work in a similar technological area.

All pairs of inventors are located on the scatterplot at Panel III according to their distances in both spaces. If both spaces were isomorphic, most of the observations would be expected to be located around the 45° line where the increment in one distance is correlated with an increment in the other one. For example, inventors *A* and *B* are closely located both in the social and knowledge space, while inventors *B* and *E* are distantly located in both of them. Both couples indicate that the closer they are in the social space, the more similar they are expected to be in functions within the firm (or the other way around). However, as explained before, when this correspondence is broken there are two possible situations according to the sign of the deviation. If social distances are smaller than knowledge distances (i.e. negative values of the mean), a cross-functional structure is observed. On the contrary, if social distances are longer than knowledge distances (i.e. positive values of the mean), it indicates the presence of redundant structures.

In the plotted example, inventors D and E are clearly a cross-functional structure since they collaborate with each other but they have very different areas of expertise. On the other extreme, inventors A and C are redundant since they work in the same area of knowledge but they do not collaborate with each other. While these cases are clearly deviated from the isomorphic line, the rest of couples may not be so obvious. In order to capture and measure deviations, for each couple of inventors the distance to the isomorphic line is calculated. This is the same of collapsing the entire knowledge-social distance space (Panel III) into the uni-dimensional space orthogonal to the isomorphic line.

Panel IV keeps only the couples of inventors into the orthogonal space to the isomorphic line and it clearly reveals the distribution of deviations. The way they agglomerate describes the structure of the organization. Many and distant deviations to one or other side of the isomorphic line would indicate the presence of redundancies or multidisciplinary structures. Panel V builds the histogram of all set of distances.

This analysis requires defining the isomorphic line. What is the proper measurement scale for each space so as to define the proportional increment in both distances that define the isomorphism? This concern is tackled throughout two different approaches. The first approach scales all distances to 1 in both spaces, such that 1 is the maximum distance registered. However, in order to avoid the influence of outliers, instead of considering maximum values as 1, percentiles 2.5 and 97.5 are considered as the boundaries of variation. Values smaller and larger than these boundaries are respectively assigned to 0 and 1 (Figure A.10 at Section A.1.3 in the Appendix illustrates the latter).

The second approach uses a different angle. Distances in both spaces are standardized by considering its mean and standard deviation. Afterwards, they are bounded to $\pm$ 3 deviations from the mean so the minimum value is given by the mean minus 3 deviations and those values smaller than that are assigned to this value. Conversely, the maximum value is given by 3 deviations from the mean and larger values are assigned to this boundary. Finally, the transformed values are scaled into the unity.

Both approaches scatter observations into the unity square where one axes corresponds to social distances and the other one to knowledge distances. The difference between both approaches lies on the scaling of observed distances (Panel IV and V at Figure A.10). Once scattering all couples of distances in this space and projecting them into the uni-dimensional space orthogonal to the isomorphic line, deviations are calculated from the isomorphism. The zero deviation stands from those couple of inventors that lie on the isomorphic line as AB in Figure 3.4. As couples deviate further, these values increase in absolute terms but with different sign depending which side it deviates to. Deviations towards redundancy are positive while deviations towards cross-functionality are negative. The whole range of variation is scaled into $[-1, 1]$ in order to simplify the interpretation, being -1 the maximum possible value of cross-functional observations, and 1 the maximum possible level of redundancy. With this information two measures are proposed to capture the type of structures predominates in the firm. One measure uses the information generated by the first approach, the other one from the second.

From each approach a different explanatory variable will be calculated. When distances are transformed according to the first approach (the one that considers the 95% of variation), the mean of the distribution will be used as explanatory variable. If the mean is not zero, it indicates that there is a higher concentration of observation at one side of the isomorphic line. The sign and the value of the mean will indicate in which area of the space, either the redundancy or the cross-functionality, the firm is mostly located. This variable will be called *mean deviation*.

When the second approach is used (the one that standardizes distances and considers only 3 standard deviations), the mean of the distribution cannot be used since it is forced to be zero. Instead, in this case the skewness of the distribution is considered. The measure is as simple as the difference between the mean and the median divided by the standard deviation. I will call it *skewness of deviations*. Positive values indicate positive skewness and vice versa. The magnitude also provides information about how much the organization deviates from the isomorphism.

3.3.3 Statistical analysis

The main goal of the empirical analysis is to explain the success of patented technological inventions (unit of observation) according to their level of interdependence and the structure of the firm which created them. The two measures previously developed, *mean deviation* and *skewness of deviations*, provide the main explanatory variables of the capacity of the firm when facing technologies with different levels of interdependence. The level of interdependence of patents is also included as explanatory variable as well as the interaction with both measures of redundancy.

The dependent variable of the statistical analysis is the number of citations. Specifically, for each patent the number of citations received during the 5-years period after its granting date is considered as the measurement of technological success of the invention. Although the innovation literature recognizes a correlation between patents' citation rates and their commercial success (Hall et al., 2001), I explicitly state *technological* success rather than other. This is mainly because I analyze inventions in a evolutionary framework where, for the sake of keeping the parallelism, successful designs are those which survive and leave their legacy in their offspring. Highly cited patents success in terms of their "offspring", i.e. inventions that use the knowledge embedded in it.

In order to collect data about citations, an algorithm sequentially extracted information about citing patents and their application dates directly from the *USPTO Patent Full Text and Image Database* online. I consider those citations that take place within the 5 years after the *cited* patent is granted by considering the application date of the *citing* patents [3]. Given the length of the considered time-windows, I discarded those patents that are not old enough to account for this period of citation in the data set (the last date of citations is 2011 October).

The empirical analysis is performed by a Generalized Negative Binomial regression (GNBR). First, this model is suitable for a discrete and positive dependent variable. Second, since the whole analysis of this chapter is based on the adaptation of NK model to the technological landscape which predicts that the variation of success is greater for high values of K, a GNBR allows modeling the variance of the dependent variable.

As mentioned before, the success of patents represented by their citation rates will be mainly associated with the interdependence levels of the invention and the socio-knowledge structure of the organization. Furthermore, in other to test hypotheses 3 and 4, the previous variables will be also related with the size of organizations. This magnitude will be captured by the logarithm of the total number of people in the co-invention network. Even though inventors are a minor part of an organization, they form the collaborative structure that deal with knowledge and therefore the object of

[3]For capturing this period of citations after the patent is granted, the minimum unit was the month. The algorithm was run on August 2013 (the on-line USPTO data base may have been modified since then)

analysis. The full size of the firm will be considered as control by the considering the total number of employees. The rest of controls included in the regression analysis are listed in the following section.

3.3.4 Controls

The first set of controls come from the use of the NK model on inventions as Fleming and Sorenson (2001) adapted it. Citations are controlled in the first place by the *number of components (N)* inventions have. Circumscribed in Kauffman's model, both interdependence and number of components determine ex-ante the nature of the design landscape. When facing the same interdependence, an invention with fewer pieces to combine represents a greater challenge than one with a large numbers.

Controlling for the number of components requires creating a dummy for a special case: the innovation with only one component. Since these innovations do not combine different technological sub-classes, there is not combination and therefore the interdependence of components cannot be defined .

The number of *previous trials* is included as control as it accounts for the number of previous combinations of the exactly same set of technological sub-classes the focal patent combines. As it counts the number of inventions that have been filed before combining exactly the same technologies, this variable controls for a possible declination in technological success of invention that are basically replica of components and structure. Furthermore, since each patent represents a local peak in the technological landscape, this variable also complements the parameter K in describing the local roughness (Fleming and Sorenson, 2001).

Similarly to the number of previous trials with the same sub-classes in the entire universe of patents, following Katila and Ahuja (2002), I will control for *search depth*. When combining different technologies to innovate, the firm must explore different ways. The success of a patent, thus, would be affected by its previous attempts. As the firm learns and becomes familiar with certain technologies, the navigation of the design landscape should be easier. That is why I control, first, for *sub-classes already used by the firm* by counting, for each patent, how many of the technological sub-classes used in this patent are already used in the firm's knowledge base. Since the number of sub-classes of patents varies, the number of already used sub-classes is divided by the total one. Therefore, this measurement lies within the [0-1] interval; 0 meaning that the firm has not used the technological sub-classes during the previous 3 years, and 1 that the firm has used at least once all the sub-classes the focal patent is classified in.

The second control regarding *search depth* regards to *couplings already used by the firm*. The firm may have already not only used the sub-classes involved in a patent, but also it may have combined them. Given that a patent is a combination of those technological sub-classes it is classified in, I control for previous combination by counting how many couplings of those sub-classes has been done in the firm's knowledge base out of the total of coupling the focal patent does. For instance, if a patent couples 4 sub-classes - A B, C and D -, then, there are a total of 6 couplings involved. If the knowledge base of the firm reveals to have coupled only sub-classes A and B, then, the value of this variable is 1/6. Logically, this variable also lies within the [0-1] interval.

The number of *prior citations* is included as a measure of the localness of search. By simply counting the number of references to other patents, this variable controls for the scope of the local search and the propensity of the sector for citing (Fleming and Sorenson, 2001). Furthermore, it also captures the combinatorial problem of an invention but from another perspective. Instead of combining technological classes, patents combine knowledge in other patents.

The number of inventors involved in the patent may also affect the difficulty of navigating certain technological landscape. Two inventions with identical interdependence may differ in the number of people directly assigned to the project and, therefore, affecting the propensity to be referenced by upcoming inventions. Controlling for the *number of authors* registered in a patent is extremely pertinent when the theoretical base of the entire chapter claims that organizations arise in order to increase the ability to deal with high complex knowledge. Consequently, this variable is included by accounting the number of inventors that formally worked together developing the focal patent.

If the previous controls try to differentiate conditions that may affect the difficulty to navigate the technological landscape, other controls should be included to isolate citations for different source of variations. The rate of citations might be influenced by the nature of the technologies and inventors operating in different industries and sectors. *Technological main class*, thus, is introduced as a control. This chapter considers the top most populated 100 USPTO's technological classes where firms in the semi-conductor industry work in. Since classes cover highly differentiated technological areas, citations rates across them might differ enough so as to control for this source of variation. Legal strategies, the technical nature of the technology, tendency to patent on the sector and complexity and scale of the technology are some of the reasons for adding this control.

Even though controlling for main technological class, being classified in more classes certainly affects the probability of being cited since it exposes the patent to different technological areas. Furthermore, the number of classes also indicates the scope of the invention. I include the *number of major classes* for capturing the level of exposure and therefore, the changes in the propensity to be cited.

Since rates of citation may also change across time, I control for *granting year*. The main reason lies on the lag between the filing and granting date. When counting citations, only those patents with filing date within the 5 consecutive years after the publication date of the focal patent are accounted since I consider that patents are generated at the moment of the filing date. Because of this, citations of newer patents might be biased: there are filed patents that cite the focal one but since those are not yet published I cannot know whether these citation lies on the citation window. Furthermore, citation might be influenced by other factors that change over the years and therefore are controlled in this variable.

While the previous controls were considered for patents, controls for firms should be also included since they differ among each other not only in their knowledge base and collaboration network and these characteristics might influence the success of their patents. In the first place, I control for the resources firms directly assign to research and development (R&D). Even though it does not fully explain the intellectual capacity of an organization, team, or even a single person, it certainly influences as it mainly provides physical means, like equipment and facilities. R&D expenditures are considered as a a share of the firm's total assets. Both measurements are obtained from the Compustat Database.

Some researchers also propose controlling for the *performance* of the firm since profitable ones may not pursue innovation as hard as a less successful one. This control is incorporated considering return on assets.

The knowledge base is assumed to manifest what firms know. Following the approach of Yayavaram and Ahuja (2008), knowledge bases were considered in 3-year windows. However previous experience might also influence the success of inventions, not only because of the accumulated knowledge but also because of phenomena like reputation or visibility. For these reason, in the statistical analysis is considered the *number of years that the firm has been patenting*.

The number of inventors who work for a company must be also considered. A greater number of inventors describe the potentiality of the social network in terms of diversity and collaboration. However, the size of the inventor staff is relevant relative to the size of the knowledge base. Because of this, the analysis includes the *ratio between* *inventors and the number of classes* of the knowledge base for capturing how much human power is devoted to the different areas the firm deal with.

As the knowledge base and collaboration network are proposed to be associated with the capacity of firms of processing high complex knowledge, especially in the simplicity of its structure measured by the hierarchical decomposability of these structures, the *size of knowledge base* must be taken into account. This variable simply counts the number of different technological sub-classes a knowledge base' patents have been classified in. It indicates the variety of technological areas the firm has managed during the 3-year period windows.

Furthermore, other measurements of structural characteristics of the firm are considered. Since both the knowledge and social structures are represented by networks, for both of them the level of *clusterization*, the number of *unconnected blocks*, and the *density* are calculated for controlling different characteristics beyond its level of decomposability. For the network of inventors is also calculated the *mean intensity of ties*.

The descriptive statistics of all variables as well as the correlation among them are shown in the Appendix, at Section A.2.

3.3.5 Results

As mentioned before, two alternative measures are proposed as main explanatory variable: *mean deviation* and *skewness of deviations*. The correlation between them is -0.6694 in the entire sample, negative as expected and strong enough to corroborate the consistency of both transformations of distances[4]. Regressions using one variable or another give consistent results. For the sake of simplicity, this Section only describes and analyzes results of the GNBR using the variable *mean deviation*.

Table 3.5 shows the results of the GNBR for the number of citations, considering the level of interdependence of those technologies involved in the invention and the presence of redundancies or cross-functional structures in the firm. Four different models are run in order to test the robustness of results regarding the main explanatory variable. Model 1 includes the main explanatory variables, those variables associated with the adaptation of the NK model, and all listed controls except those regarding the structural characteristics of both the social and the knowledge network of the firm. Model 2

[4]When considering only those observations that corresponding to the largest firms (size of collaboration network superior to its median) the correlation is even stronger (-0.8234).

Figure 3.5: Generalized negative binomial estimates of citation counts (5-year window, standard errors in parentheses)

General Negative Binomial model citation counts

Variables / Models	**MEAN**				**VAR**			
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
constant	3.7271 (0.285) ***	3.3245 (0.346) ***	1.6969 (0.694) *	0.6618 (0.599)	-0.4636 (0.473)	-0.5016 (0.538)	-1.6183 (1.008)	-1.3639 (0.809) +
prior citations	0.0058 (0) ***	0.0059 (0) ***	0.0059 (0) ***	0.0076 (0) ***	-0.0003 (0)	-0.0003 (0)	-0.0002 (0)	0.0001 (0)
dummy single subclass	-0.3021 (0.135) *	-0.3082 (0.135) *	-0.31 (0.135) *	-0.2561 (0.138) +	0.2323 (0.181)	0.2196 (0.18)	0.2169 (0.181)	0.2512 (0.179)
number of major classes	0.0417 (0.007) ***	0.0418 (0.007) ***	0.0416 (0.007) ***	0.0495 (0.008) ***	-0.0013 (0.01)	-0.0009 (0.01)	-0.0012 (0.01)	0.0013 (0.01)
previous trials	-0.0002 (0)	-0.0003 (0)	-0.0003 (0)	-0.0003 (0) *	-0.0006 (0) *	-0.0006 (0) *	-0.0006 (0) *	-0.0006 (0) *
number of authors	0.0539 (0.003) ***	0.0535 (0.003) ***	0.0535 (0.003) ***	0.0492 (0.003) ***	-0.0096 (0.004) *	-0.0098 (0.004) *	-0.0099 (0.004) *	-0.0097 (0.004) *
couplings already used by the firm	-0.0225 (0.013) +	-0.0224 (0.013) +	-0.022 (0.013) +	-0.0273 (0.014) *	-0.0251 (0.018)	-0.0254 (0.018)	-0.0255 (0.018)	-0.0285 (0.018)
sub-classes already used by the firm	0.1309 (0.016) ***	0.1325 (0.016) ***	0.1319 (0.016) ***	0.1254 (0.017) ***	-0.194 (0.023) ***	-0.1914 (0.023) ***	-0.1921 (0.023) ***	-0.1875 (0.023) ***
inventors by class	0.1007 (0.014) ***	0.1021 (0.022) ***	0.0968 (0.022) ***	-0.0945 (0.017) ***	0.0003 (0) *	0.0007 (0) **	0.0007 (0) **	0.0003 (0) *
age	0.0006 (0.006)	-0.0029 (0.006)	-0.0017 (0.006)	0.0166 (0.002) ***	0.0296 (0.007) ***	0.0276 (0.008) ***	0.0288 (0.008) ***	-0.0135 (0.002) ***
R&D intensity	-0.6388 (0.266) *	-0.4509 (0.279)	-0.4706 (0.281) +	0.0641 (0.133)	-0.542 (0.379)	-0.798 (0.392) *	-0.8048 (0.394) *	-0.747 (0.182) ***
Number of employees (log)	-0.0768 (0.019) ***	-0.0611 (0.02) **	-0.0618 (0.02) **	-0.0456 (0.011) ***	0.0891 (0.028) **	0.0836 (0.029) **	0.0789 (0.03) **	-0.039 (0.015) **
Performance	0.0008 (0) +	0.0008 (0) +	0.0007 (0) +	0.0021 (0) ***	-0.0003 (0.001)	-0.0004 (0.001)	-0.0004 (0.001)	-0.0012 (0.001) *
1/N	-1.3306 (0.208) ***	-1.3411 (0.208) ***	-1.3432 (0.208) ***	-1.2096 (0.214) ***	0.7369 (0.274) **	0.7191 (0.273) **	0.7148 (0.274) **	0.7965 (0.273) **
1/N squared	1.1824 (0.309) ***	1.1952 (0.309) ***	1.1985 (0.309) ***	1.0327 (0.317) **	-0.7964 (0.41) +	-0.7711 (0.409) +	-0.7666 (0.41) +	-0.8481 (0.407) *
Interdependence (K)	-0.6763 (0.185) ***	-0.6431 (0.185) **	-0.6389 (0.184) **	-0.8917 (0.189) ***	-0.2058 (0.262)	-0.1986 (0.263)	-0.1989 (0.263)	-0.4739 (0.261) +
K squared	0.1715 (0.043) ***	0.1658 (0.043) ***	0.1657 (0.043) ***	0.1947 (0.048) ***	-0.4513 (0.165) **	-0.4711 (0.171) **	-0.4854 (0.175) **	-0.3324 (0.173) +
K/N	0.2305 (0.183)	0.2341 (0.183)	0.2336 (0.182)	0.341 (0.191) +	0.736 (0.272) **	0.7453 (0.272) **	0.7531 (0.273) **	0.7525 (0.274) **
mean deviation	-1.6532 (0.402) ***	-1.4407 (0.418) **	-0.6172 (0.535)	1.0175 (0.432) *	0.7125 (0.606)	0.6266 (0.614)	1.2479 (0.765)	0.6946 (0.561)
mean deviation squared	-0.3031 (0.223)	-0.5333 (0.23) *	-0.3929 (0.246)	-0.7377 (0.208) ***	-0.47 (0.324)	-0.3513 (0.335)	-0.3706 (0.368)	-0.1625 (0.297)
mean deviation x K	**0.7581 (0.27)** **	**0.6525 (0.271)** *	**0.6469 (0.27)** *	**0.8412 (0.271)** **	0.3781 (0.394)	0.3542 (0.396)	0.3632 (0.396)	0.5856 (0.385)
Log number of inventors - size	-0.3509 (0.03) ***	-0.3173 (0.046) ***	-0.0909 (0.102)	0.0692 (0.094)	0.0751 (0.051)	0.0344 (0.062)	0.2084 (0.142)	0.2774 (0.124) *
mean deviation x size	0.2635 (0.056) ***	0.2578 (0.059) ***	0.1191 (0.082)	-0.1362 (0.07) +	-0.0498 (0.085)	-0.0449 (0.086)	-0.1495 (0.115)	-0.0783 (0.089)
KB's Clusterization		-0.2216 (0.187)	-0.1679 (0.188)	0.5219 (0.154) **		0.0116 (0.275)	0.0351 (0.277)	-0.1855 (0.204)
KB's Density		2.3186 (1.581)	2.807 (1.595) +	4.7252 (1.06) ***		7.9706 (2.06) ***	8.3059 (2.066) ***	-2.2533 (1.117) *
KB's blocks		0.0028 (0.001) *	0.0031 (0.001) *	-0.0025 (0.001) *		0.0009 (0.002)	0.0012 (0.002)	-0.001 (0.001)
SN's blocks		-0.0003 (0) +	-0.0001 (0)	0.0007 (0) ***		0 (0)	0.0002 (0)	0.0001 (0)
SN's density		-1.1876 (2.729)	-2.2634 (2.777)	1.0156 (1.975)		-7.3632 (3.847) +	-8.4245 (3.905) *	6.4973 (2.463) **
SN's clusterization		0.5615 (0.181) **	0.5661 (0.181) **	0.201 (0.114) +		0.3751 (0.267)	0.3714 (0.267)	-0.4369 (0.148) **
SN's mean tie		-0.004 (0.019)	-0.0093 (0.02)	-0.0417 (0.015) **		-0.0513 (0.026) *	-0.0562 (0.026) *	-0.01 (0.021)
standard deviation of deviations			3.8465 (1.458) **	1.7647 (1.334)			2.8548 (2.137)	3.2448 (1.826) +
standard deviation of deviations x K			-0.5784 (0.244) *	-0.1248 (0.223)			-0.4769 (0.352)	-0.4238 (0.304)
Dummy Controls for main technological class / firm / granting year				No dummy for firm	Dummy for main technological class / firm / granting year			No dummy for firm

Num observations	85181	85181	85181	85181
Prob > chi2	0.0000	0.0000	0.0000	0.0000
Log pseudolikelihood	-273888.72	-273856.54	-275373.24	-273849.66

(Standard Error) *** p<0.001 ** p<0.01 *p<0.05 +p<0.1

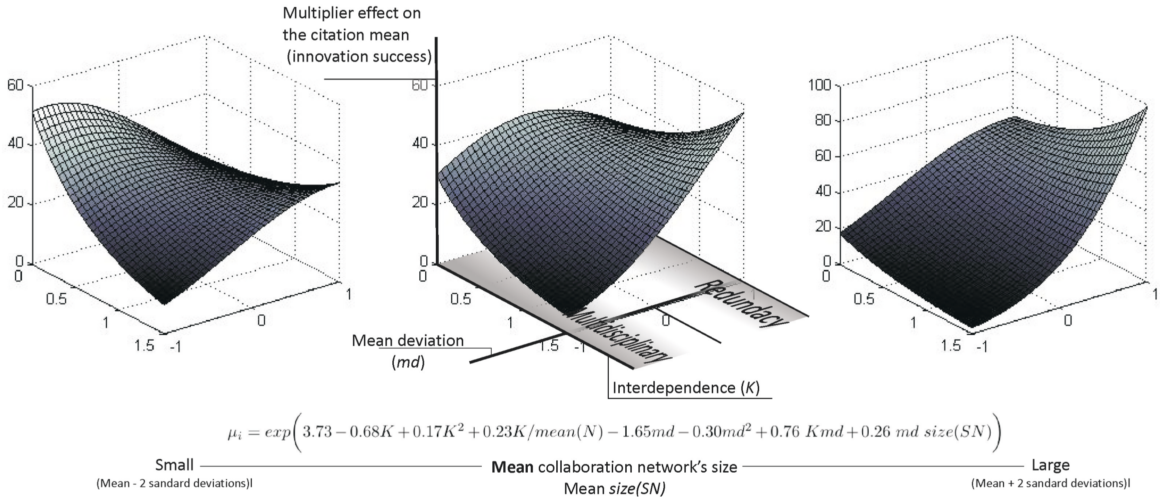


Figure 3.6: Expected mean multiplier effect as a function of mean deviation (*md*) and interdependence (*K*) estimated from Model 1.

includes standard measurements of network characteristics that describe the structure of both the firm's social and knowledge network so they can compete with main explanatory variable (which also describes both structures) for explaining the variation of citations. Model 3 includes two controls that explain the dispersion of the distribution of deviations regarding the isomorphism: the standard deviation and the standard deviation interacting with the interdependence parameter. This is done in order to control for a possible displacement of the mean due to a high dispersion rather than a general deviation to redundant or cross-functional structures. Finally, Model 4 discards the dummy variable for firm given that many controls account for their individual characteristics.

Results regarding the main variable that test both Hypotheses 1 and 2, *mean deviation x K* are significant for all models, keeping the sign and similar values. The interpretation of the value of estimated coefficients is difficult since the construction of the measure has many transformations. Since there are many interactions, Figure 3.6 helps reading Table 3.5. The interaction term between interdependence of innovations' components and mean deviation presents a positive estimated coefficient in all models. This clearly indicates that while the interdependence of technology increases, redundancies increment the success of inventions, and oppositely, when interdependence decreases, cross-functional structures perform better.

If an organization which deals with medium interdependent technologies has a collaborative structure that mirrors the knowledge structure (isomorphic) but increases in two standard deviations the level of redundancy (variable *mean deviation*) in its medium sized collaborative network, the mean expected citation rates of those patents produced increases in nearly 5%. Oppositely, if the organization shifts to multidisciplinar structures in two deviations, the expected mean of citation decreases in 11.6%. However, when the firm faces more extreme problems, the collaborative structure acts differently. When dealing with very low interdependent technologies, the effect of deviating from the isomorphism between the knowledge and the social structure are negative in each direction. When tilting to multidisciplinar structures in two deviations, the expected success decreases in 7% and 0.5% when shifting to redundant structures. On the other extreme, when dealing with highly interdependent technologies, the expected success of inventions increases in 10.5% when generating redudancies and strongly decreases in 16% when shifting to multidisciplinar structures.

The effects on expected success of inventions exaggerates when the size of the firm is considered. As Figure 3.6 and Figure 3.7 shows, small firms perform better with low interdependent technologies and multidisciplinar structures. When dealing with low interdependent technologies, generating muldisciplinar structures outperforms isomorphic ones by 9.7%. On the contrary, if the organizations has redundant structures, the performance is 15.5% inferior than the isomorphic case (23% lower than the multidisciplinar structure). When firm have large networks of inventors (mean size plus two deviations), redundant structures outperform multidisciplinar ones in any type of technology.

Figure 3.6 plots this idea in the central panel. When the structure of an organization merges diverse knowledge throughout the social space, it performs better in smooth technological landscapes as they require disruptive thinking to find successful innovations. However, as the interdependence of technology increases, those same structures badly perform. On the other extreme, when organizations present redundant structures are better prepared for high complex scenarios where they can find successful designs without clues or patterns helping.

Since the empirical analysis include an interaction term between the size of the collaboration network within the firm and the mean deviation from the isomorphism, Figure 3.6 includes two more panels. The left panel shows the estimated effects when the social network's size is equal to the mean of the data set minus two standard deviations, while the right panel shows the mean plus two standard deviations of the size distribution. The interaction term is included in order to analyze the role of redundancies according to different scales of networks of inventors. As it can be observed, for small networks the effects of redundancies and cross-functional structures are exaggerated in support of both Hypotheses 1 and 2. In low interdependent scenarios where disruptive

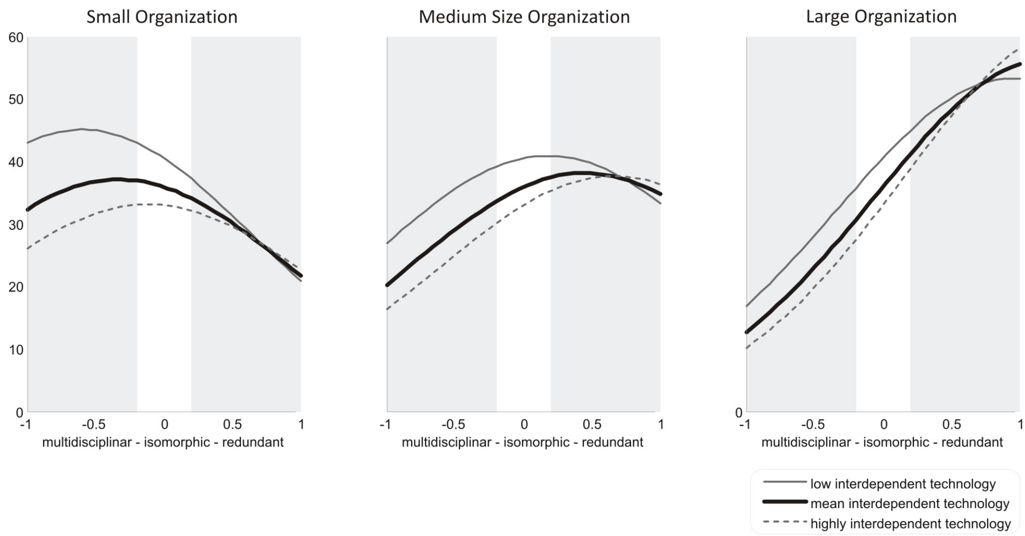


Figure 3.7: Expected mean multiplier effect as a function of mean deviation (*md*), plotted by 3 different levels of interdependence (K) and size of the collaborative network.

thinking is highly needed in order to find new successful combinations of knowledge, cross-functional structures clearly perform better than redundant ones. This relation is inverted for high levels of interdependence. On the other hand, for large networks of inventors, even though the relation between complexity and redundancy or cross-functional keeps the coefficient's sign, redundant structures seems to dominate the entire scenario. This may be due to the impossibility of avoiding redundancies when the network is large enough.

Section A.3 depicts estimated coefficients for a regression analysis changing the variable *mean deviation* for *skewness of deviations*. As mentioned before, these two variables are negatively correlated. By construction, negative values of *skewness of deviations* indicate multidisciplinar structure while positive values indicate redundant structures. Moreover, similarly to the metric *mean deviation*, the magnitude of *skewness of deviations* describe the *level* of redundancy or cross-functionality. Larger values depict larger dissimilarities between the knowledge structure and the social structure. Table A.14 shows the results of the regression analysis and Figure A.13 plots them the same way Figure 3.6 did it. The relations between interdependence, structure (redundancy and cross-functionality)and size are similar to the ones analyzed before. In other words, the empirical analysis support all four hypotheses. More importantly, these results supports the metric rather than the hypotheses.

A final note should be included. A generalized negative binomial model was used for analyzing the relation between structure, interdependence of technology and size of the firm. In order to do this, several controls were used as dummy variables including firm and granting year of the patent. Instead of following this strategy, I could have used a negative binomial regression with fixed effects on granting years and patents. As it was checked, the empirical results of using this model give the same sign and similar magnitudes for the estimated coefficients of the main variables. However, this model had to be discarded since it does not allow modeling the variance of the dependent variable. Differently to the Poisson model, the negative binomial regression allows the variance being different to the mean of the dependent variable. However, it is assumed to be constant. Since the whole analysis of this chapter is based on the adaptation of the NK model to the technological field, the variance cannot be assumed constant across the range of variation of K. Assuming the opposite invalidates the whole model. Because of this reason, a negative binomial regression with fixed effects was not explicitly included.

3.4 Discussion

One of the main contribution of this chapter lies in the operationalization of the concept of structural redundancy in a continuum that has cross-functionality as opposite. This way, the proposed operationalization incorporates a well known team or organizational structure, cross-functionality, as it also defines and captures redundant structures. Furthermore, this approach questions a generally and implicitly assumed isomorphism between social and knowledge spaces across the dominant literature on innovation and social networks Chapter 2. By considering social ties as channels where knowledge flows through, social proximity is usually assumed to describe knowledge proximity. Them, by unfolding the assumed isomorphism into the two spaces involved, not only misunderstandings are avoided but we can analyze what happens when organizational structures deviate from this correspondence.

The operationalization of these two concepts, organizational redundancy and cross-functionality, may also embrace many distinct topics in the innovation literature in a coherent framework. As mentioned before, concepts as knowledge overlapping, slack, geographical dispersion, weak ties (Granovetter, 1973), structural holes (Burt, 1992), simmelian ties (Tortoriello and Krackhardt, 2010), clusters of firms (Bell, 2005), scientific collaboration (Owen-Smith and Powell, 2004b; Newman, 2001), and the role generalist in innovations melero2015renaissance among others. This framework could be also expanded by relating these general structures, redundant and cross-functional, with organizational hierarchies.

The empirical analysis focuses on innovation activity, particularly in patented innovations in semi-conductor industry. The information used to study the social structure is co-invention, and the one to study the interdependence of inventions are technological classes. Although this approach has been used many times before in the innovation literature, it is not free from limitations. Some of them were mentioned before, especially regarding co-authorship. However, the main concern of using this approach would be the possible correlation between co-authorship and expertise proximity. Two inventors co-authoring a patent, by this operationalization, are socially tied as well as close in the knowledge space. In other words, the way the social space and the knowledge space are observed would force the isomorphism. However, as it can be observed in Section A.2, the variation of distances is large enough to provide a large variation of the metric. Furthermore, in order not to exclusively rely on the main metric, i.e. *mean deviation*, another one is built and tested providing coherent results. These are shown in the Appendix at Section A.3. A third metric was designed but rejected before using it for empirical analysis as it showed a high correlation with *mean deviation*. Anyways, it served to confirm the realiability of the *mean deviation* given its different methodology. Equation A.22 at the Appendix shows this measure and its analysis.

As results suggest, redundant structures overcome the complexity of dealing with highly interdependent technologies. Fleming and Sorenson (2001) provided empirical evidence for supporting the idea that "the level of interdependence among innovations' components describes the difficulty for perfecting as well as it indicates the possibility of some very useful configurations". Specifically, they found that high levels of interdependence were associated with a expected decrease in innovation's success as well as with an increment in the variance. Using their proposed measure for capturing interdependence, this chapter shows that higher levels of interdependence are not necessarily associated with lower expected values of success. The empirical analysis suggests that those very useful though rare configurations spread across rugged landscapes can be found when organizations use parallel exploration.

Fleming and Sorenson (2001) also proposed that low levels of interdependence would not allow successful innovations. Differently to the previous case, the reason behind this statement is the the lack of altitude of those peaks in smooth landscapes. This chapter shows coherent results since expected success in these landscapes is not larger than in rugged landscapes in any case. However, this chapter suggests that expected success on smooth landscapes depends on how firms process knowledge. Multidisciplinary or cross-functional structures are better than redundant configurations finding peaks by jumping across large basins of attraction. However, this effect depends on the size of the firm.

This chapter also suggests that small firm perform better when they rapidly expand the knowledge area covered by their members by forming multidisciplinary groups and by thinking out of the box. These results might be describing "garage startups" as Apple or Hewlett-Packard when they were founded: a radical idea, small size and a multidisciplinary and close team. Redundancy, as the way its operationalized here, may be not so much as undesirable as impossible in these firms. In order to exist, redundant components must operate almost independently, and when the organization is really small, there is not such a thing.

On the contrary, redundancies might be unavoidable in large firms due to the expansion of the social space. The connectivity capacity of people is not infinite but quite the opposite (Dunbar, 1992; Gonçalves et al., 2011), and if the size of a network considerably increases, socially distances irremediably increase. Therefore, specialists in similar areas will end up working in parallel. Redundancies may naturally emerge when firms grow large or geographically expand, when they endure many decades, or when they acquire other firms. The idea of efficiency may lead managers to dedicate great efforts in combating and reengineer the existence of overlapping processes, departments, systems or functions. However, as this chapter suggests, redundancies may play a crucial role. What it seems noisy and disordered, it may be indeed an efficient machinery that process knowledge in ways we cannot fully understand. Simon (1962) suggested that "networks poised at the edge of chaos can perform the most complex tasks".

Although this chapter focuses on innovation, the logic may apply to broader areas. The concept of interdependence exceeds the field of biology. Kauffman (1993) proposed the NKmodel for depicting the interaction of genomes and its consequences on evolutionary processes[5]. However, as he also recognizes, this logic goes beyond genetics.

[5]A good example to this chapter's hypotheses is biological evolution itself. The cross-functional tactic, the one that involves gathering distinct knowledge in the same team in order to increase the distance of jumps, is represented by sexual reproduction. It has been postulated that the emergence of gender and sexual reproduction of leaving beings accelerate evolution by mixing two different genomes (Birdsell and Wills, 2003) and therefore increasing the design variation of offspring. On the other hand, the redundancy tactic, the one that involves assigning different parts of the system to perform the same function, is represented by reproductive isolation (Mayr et al., 1963). This would be the case of groups of organisms that live long periods without mixing with other groups of the same species. These groups would be redundant in the sense that they are making evolve the same species. If they never mix, they could end up generating

Any system with interacting components can be described in terms of this model in order to capture the difficulty of finding the best configurations. Simon (1962) described many complex systems that were formed (or perceived as formed) by many interacting components sorted in nearly decomposable layers arranged in a multi-layered architecture, from linguistics systems to organizations themselves. And wherever there is a system with many parts that interact in a non trivial way, interdependence can be applied to its analysis. Therefore, the generality of this logic can be extended to problems beyond innovation. For instance, Levinthal (1997) studied organizational adaptation to changing environments by using the NK model. Other researchers that use the logic of the NK model for describing all kind of problems organizations may deal with suggest that perhaps the role of redundancies when facing highly interdependent problems extend beyond innovation. Of course, empirical results in this chapter as circumscribed to the semi-conductor industry.

This chapter opens the question: how do redundant structures affect the firm performance and especially its cognitive capacity? Parallel processing are known to be a crucial characteristic of human brains: the neural network which constitutes it is full of redundancies. Indeed, this is postulated to be source of our ability to induce patterns and being creative (Heilman et al., 2003; Haykin, 1999). These architectures are being used as inspiration for writing the most sophisticated algorithms in computation, from patter recognition to learning programs. Much time has passed since the ideas of scientific management of Weber and Taylor that sought to accurately design every piece of organizational processes. Nowadays it is recognized that managers do not have and do not seek total control on organizations, especially those who are focused in knowledge-intensive sectors. The formal structure is not entirely manageable since much information uses channels outside the formal chart generating a complex system. This chapter, then, contributes to this idea. Something as easily called inefficient as redundant parts may play a substantial role in enhancing the adaptative performance of the organization, not only by providing reliability but also by fostering its cognitive capacity for dealing with high interdependent problems. This is the reason to this chapter for emphasizing redundancy instead of cross-functionality. Cross-functionality seems much easier to accept for managers than redundancy. Its effects are more intuitive and not necessarily condemned as inefficient, as well as they have been previously studied in the innovation literature. Oppositely, arguing that organizational redundancies may contribute to the performance requires much more rationalization.

Future research could dig deeper in those mechanisms that make redundancies out-

different species.

perform cross-functionality with high interdependent technologies, and cross-functionality outperform redundancies with low interdependent ones. Theoretical models could be used for simulating and testing different mechanisms. Moreover, this could be extended by relating organizational redundancies and cross-functionality with other outcomes or phenomena rather than innovation.

"It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience" (Einstein, 1934)

Chapter 4

The Intelligence of Organizations

The Challenge of Innovating by Decomposing Complexity

Abstract

This chapter proposes using innovations as a way of observing the intellectual capacity of organizations. Innovations provide both a measurement of success and a measurement of intellectual difficulty. Using Kauffman (1993)'s understanding, the difficulty is explained in terms of the complexity of finding suitable configurations of interdependent components. When the success of different innovations is compared by controlling for their difficulty, the intellectual capacity of organizations is revealed. This chapter, then, explains this capacity in terms of the organizational ability of simplifying the complexity into hierarchical decomposable structures. For that purpose, both the firm's knowledge base and collaborative network among inventors are analyzed in terms of Simon (1962)'s conception of near-decomposability. The statistical analysis of patented inventions from firms operating in the Semi-conductor industry supports expectations: the ability of decomposing complexity heavily influences the intellectual capacity of organizations when innovating with highly interdependent technologies.

4.1 Introduction

Organizations are sophisticated social machineries that process vast volumes of information and knowledge on daily bases. Providing a product or service can represent an extremely complex challenge, and as such, it requires numerous people working together. By facilitating the collective work, organizations generate an intelligence capable of solving a wide variety of overwhelmingly complex problems (Glynn, 1996). Among them, a pure intellectual problem stands out from the rest: innovating. When creating new technologies, organizations work at the very knowledge frontier and their intellectual capacities are used to their maximum. Innovations, then, can reveal an interesting dimension of an organization's ability to deal with complex problems. This defining characteristic of firms has almost no empirical attempts of measurement despite the interesting implications that could have for management. In the following pages I propose a way of explaining the capacity to innovate in terms of the intellectual capacity of an organization as a whole.

This paper understands organizations as social arrangements that generate an intellectual machinery capable of innovating in complex technologies. Based on the adaptation of Kauffman's NK model (Kauffman and Levin, 1987) to the technological field by Fleming and Sorenson (2001), the interdependence of an innovation's components is considered to describe the technical difficulty of achieving a successful configuration. *Ceteris paribus*, the relation between technical difficulty and success of innovations would reveal the organizations' underlying intellectual capacity.

The goal of this paper is to explore how organizations in a given industry succeed when innovating with technologies of different complexity. This revealed ability is assumed to manifest the firm's intellectual capacity. As a collective characteristic, it is explained by the firm's social structure. As intellectual, it is explained by the firm's knowledge base. Specifically, the firm's intellectual capacity is related to the degree of hierarchical decomposability of the firm's social structure since this accounts for its adaptability and capacity for information processing (Galbraith, 1977); and it is related to the degree of hierarchical decomposability of the knowledge structure since this reveals the ability for simplifying complex phenomena, for recognizing patterns, for understanding and therefore for solving problems.

According to this perspective, the paper conducts two analyses. First, an explicit measurement of the intellectual capacity of firms is proposed by combining both the success and the complexity of innovations. This measurement is explained according to the levels of decomposability of the dual-structure of firms. Afterwards, as a robustness analysis for the impact of the explanatory variables on the success of inventions, the measurement of intellectual capacity is left out and the success of innovations is explained in terms of the complexity of the technology involved and the levels of decomposability of organizations. Empirical results are consistent and show that the complexity of technologies is not necessarily associated with less successful innovations on average but it depends on the intellectual capacity of the organization.

4.2 Theory and hypotheses

Among the different theories that explain the existence of organizations, some of them focus on their capacity for increasing the collective intelligence of a group as it reduces the knowledge controlled by its members (Kogut and Zander, 1996; Adler and Kwon, 2002; Kalkan, 2011). Throughout history, the constant growth of our collective capacity for learning and knowing has not been related to an increase in our brain's capacities but to the advancement of our social complexity. The human intelligence has been a social phenomenon since its origins, from the moment first humans gathered in communities for survival to the current global society

(Eisenberg, 1981; Ridley and Wolf, 1998). As the complexity of knowledge grew along with the development of societies, organizations emerged as the main intellectual engines of this process.

Organizations play a major role in the development of knowledge and technology mainly because of its capacity of enhancing the collective intellectual activity (Hargadon and Sutton, 1997; Gabbay and Zuckerman, 1998; Tsai and Ghoshal, 1998; Nahapiet and Ghoshal, 1998; Guimera et al., 2005). As knowledge and technology develop, collaboration has become utterly necessary for researching and innovating. However, collaboration is not easy to achieve. Its backbone is exchange of knowledge (Nahapiet and Ghoshal, 1998; Cross et al., 2001; Phelps et al., 2012) and sometimes, due to the characteristics of the content or the persons involved, this might be extremely difficult[1], especially when operating at the knowledge frontier. Thus, the social processing of information is not only a matter of gathering large groups of people but it also depends on how strong collaboration ties are. This might happen only within organizations as they boost cooperation and knowledge transmission.

The key of this capacity lies on the meta-identity that an organization develops. As Kogut and Zander (1996) proposed, organizations provide a context of discourse where "coordination, communication, and learning are situated not only physically in locality, but also mentally in an identity". Large firms may employ thousands of persons, each of whom being only capable of dealing with small pieces of much larger problems, but the firm's over-arching narrative articulates members so they work like a unique intellect (Stephenson, 2011). Organizations, hence, can be understood as social organisms whose intellectual capabilities are largely superior than the sum of its members' and whose complexity might be so vast that the individual contribution becomes difficult to distinguish from the overall organizational intelligence (Nonaka and Takeuchi, 1995, 1997).

Although not always explicitly, the concept of organizations as a collective intelligence has a long record in the management and sociology literature (Glynn, 1996; Hayek, 1945; Kalkan, 2011; March, 1991; Walsh and Ungson, 1991; Weick and Roberts, 1993; Williams and Sternberg, 1988; Woolley and Fuchs, 2011). However, few attempts have been made of empirically capturing this very intuitive and potentially useful idea. Different measurements for a firm's performance have been used in the innovation literature, such as financial information, market indicators, or commercial success, while none of them capture its ability for processing knowledge (March and Sutton, 1997; Nonaka, 1994).

This research uses innovations as a way of looking into the intellectual capacity of organizations. Technological innovation offers the perfect ground because of its intellectual nature (Glynn, 1996) and because of the difficulty that implies. In the first place, organizations might be the only ones capable of innovating in complex technological areas. The widening gap between an average person's intelligence and the intrinsic complexity of the knowledge we are immersed in, makes innovating an enterprise that can be rarely undertaken by a single or few individuals

[1]There are many researchers that has proven the difficulty of knowledge to be transmitted, for instance Jaffe et al. (1992); Almeida and Kogut (1999)

(Singh et al., 2010). Consequently, innovating at most of fronts of technology requires such a degree of coordination and intensity of collaboration that is only achievable by organizations.

In the second place, the very nature of innovation is intellectual. Organizations actively exercise their cognitive apparatus when generating new knowledge and technologies. They must explore, interpret, learn, diagnose, master and combine knowledge (Schumpeter, 1939; Walsh and Ungson, 1991; Glynn, 1996). Working at the technological frontier implies not only dealing with advanced knowledge (and therefore difficult to acquire), but it also implies constantly facing unknown technical problems, having no references to consult or sometimes not even having a proper language to describe the phenomena. Furthermore, differently to other intellectual activities, innovating in high technological industries works under market pressure: problems must also be solved faster and more efficiently than competitors. Innovating, thus, combines the challenges of learning the most advanced knowledge, the newness of unknown intellectual territory, and the constant pressure from competition. As an intellectual activity, it provides a valuable manifestation of the cognitive abilities of organizations as it demands their very best.

Innovating can be regarded as a test for the intellectual capacity of organizations. In the first place, it is a test of the capacity of the entire firm rather than its individual members. Organizations form such complex networks of interactions actively transmitting information and knowledge that any of their actions should be though of as a product of the whole (Kogut and Zander, 1992; Radner, 1993; Bolton and Dewatripont, 1994; Kogut and Zander, 1996; DeCanio and Watkins, 1998). Exploring the technological field for new ways of solving problems is not the exception but its maximum consummation. Innovations, thus, depict the firm rather than the individual.

In the second place, when patented, innovations can work as a test since they present both elements for assessing the intellectual capacity: a measurement of success and a measurement of difficulty. Regarding the first, the commercial success of patents would indicate functioning, efficient and successful solutions for market needs. On the other hand, defining the difficulty of innovating requires more sophistication. This paper understands difficulty as the intellectual challenge is posed by the technological *complexity* of an innovation's components. This interpretation is based on the application of NK model by Fleming and Sorenson (2001) to the innovation field. It essentially conceives complexity as a roughed landscape where solutions are its peaks and failures are its valleys (Kauffman and Levin, 1987). The technological landscape is tuned by two parameters: N, the number of components of an innovation, and K, the level of interdependence of those components (how much they affect each other in their individual performance). As the interdependence of components increases, the landscape become rougher and it is difficult to find the best peaks (Figure 4.1). In other words, for high levels of complexity, inventions are expected to fail more on average while successful configurations become rarer but more successful than in smooth landscapes (Fleming and Sorenson, 2001).

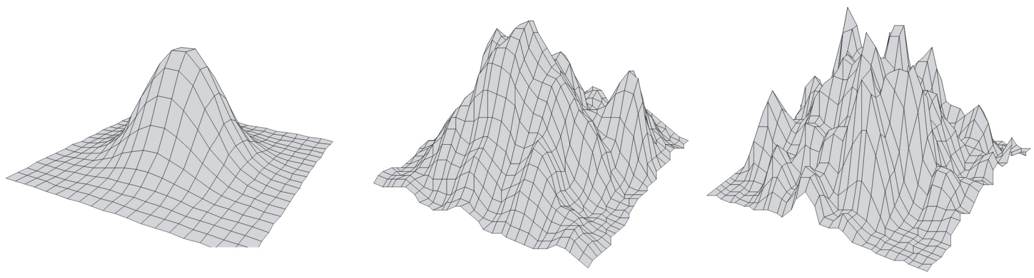


Figure 4.1: Different technological landscapes according to different values of the interdependence parameter K

By considering the complexity of an invention and its success, it is possible to observe the firm's ability to deal with complex problems, namely, their intellectual capacity. A high level of complexity constraints the likelihood of success for a blind searcher of configurations[2], but at the same time it generates few but great opportunities for a sophisticated intellect. Organizations do not try every possible combination but actively discard and select technological paths before even trying them. Thus, when innovating in overwhelmingly complicated scenarios, the capacity of succeeding depends on the ability to map the complexity into a simplified form that preserves its main patterns. The organization's intellectual capacity would be related to the ability of *decomposing* the complexity.

As proposed by Simon (1962), a cognitive strategy for untangling complex phenomena consists in decomposing the whole into a simplified form that captures the main patterns. The human intelligence is intrinsically associated with its capacity for pattern recognition (Kurzweil, 2012). The ability to simplify complexity it is *understanding* itself. According to (Einstein, 1934), "it can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience". Embracing complexity does not mean difficulty: it is up to the intelligence that deals with it. And if sufficiently powerful, complexity offers the maximum simplicity (Berlow et al., 2009).

Because of the intrinsic relation between intelligence and simplification, Simon's approach is used to explain the intellectual capacity of organizations. Specifically, Simon proposed that complex systems can be perceived, comprehended and understood

[2]A blind searcher is an evolutionary algorithm, a criterion that randomly changes a design and evaluates it. According to the evaluation, the blind searcher keeps it or discard it, and the process repeats itself. It does not pursuit a goal but mutates randomly and marginally.

whenever their patterns of interactions form a hierarchical near-decomposable (HND) architecture. This structure is characterized by many layers of clusterization, where in each layer elements with strong interaction agglomerate and the rest of interactions are almost negligible. If a system is not HND, our brain is neither able to process nor to perceive it. In words of Cohen and Levinthal (1990) when explaining the capacity of absorbing knowledge by firms: "the breadth of categories into which prior knowledge is organized, the differentiation of those categories, and the linkages across them permit individuals to make sense of and, in turn, acquire new knowledge" (Bower and Hilgard (1981) as cited by Cohen and Levinthal (1990)). Rather than an ontological commitment about the nature of the universe, understanding phenomena as HND is more a cognitive strategy for simplifying a problem to a level we can understand. Consequently, as organizations navigate the vast complexity of the technological landscape, their intellectual capacity to tackle such scenario should be strongly associated with how much they are able to decompose the landscape into a simplified map.

The ability of decomposing complexity can be observed on how organizations structure themselves. If we think of information-processing machines we will think on brains and computers. Both of them have two intrinsic dimensions: the conceptual structure and the physical support, the cognitive map and the neural network, the software and the hardware, or the epistemological and ontological dimensions in terms of Nonaka (1994). As information-processing machines, organizations also have this double nature: a knowledge base which depicts their cognitive map and a collaboration network that physically support the latter.

Regarding the first, the understanding and interpretation of complexity can be observed on how organizations structure their knowledge (Yayavaram and Ahuja, 2008). When patented, innovations not only offer a way for observing the intellectual ability of firms by describing the complexity of problems and the success of those solutions, but they also provide information about *what* organizations know and *how* they combine and relate knowledge. It draws the structure a firm's cognitive map of the technological landscape, and therefore, it allows observing how the firm decompose the technological complexity.

The knowledge base plays a major role in innovating as it influences and is influenced by the research activity of the organization (Cohen and Levinthal, 1990). The problem-solving processes that involve trial and error are proportionally to the difficulty and novelty of the problem. This process is not random but highly selective based on a criterion of progress to the goal (Simon and Ando, 1961). When the nature of the problem is unknown because of its newness (as it happens when innovating) the criterion of progress to the goal is not objective but quite the opposite. The knowledge base, then, is the map of the technological landscape that orientates firms' inventors on their navigation in the search of peaks. As it is related to the organizational structure, beliefs, routines (Nelson and Winter, 2009), organizational memory (Walsh and Ungson, 1991; Moorman and Miner, 1998) and communication patterns (Allen, 1984a) among others, the knowledge base works as the heuristic that guides the research activity of a firm. Therefore, as the landscape become more complex, the ability of the firm to decompose this complexity into a simplified version should determine the capacity to find successful innovations.

Hypothesis 5. *A firm's intellectual capacity will be positively related to the degree of hierarchical decomposability of its knowledge base.*

The simplification of complexity throughout a hierarchically decomposable structure plays a major role in understanding a phenomenon the more complex it is. As firms innovate, learn, grow and expand in their knowledge bases, the degree of decomposability of these structures becomes more relevant for its success. While the decomposability of a start-up firm that is narrowly focused on specific technology would not make the difference, it will matter for large firms dealing with many technological fronts, years of experience and quite varied knowledge backgrounds of their employees. Then, it is proposed the following hypothesis.

Hypothesis 6. *The relation between the intellectual capacity of a firm and the degree of hierarchical decomposability of its knowledge base will be stronger as the knowledge base is larger.*

When it regards to the physical support of the knowledge base, organizations can be also analyzed in their internal social structure. The innovative activity of a firm not only depicts the way it knows and understands technology, but it can also describe the collaborative structure within. As it was said before, one of the reasons organizations play a major role in innovating lies on their capacity of gathering people and enhancing the collaborative work. This way organizations are able to innovate in much more complex fields than an individual could.

Even though it could be said that nowadays there is not such a thing as a lonely inventor because of the social influence there is in any intellectual activity of a person (Granovetter, 1973; Phelps et al., 2012), organizations push collaboration to a different level. Since the knowledge involved in an innovation can easily outrun a single person's mental capacity, the ability for transmitting it among people is crucial. However, this can be very difficult to achieve: sometimes knowledge might be tacit (not codified), others it may require a shared background to be transmitted (Ethiraj and Levinthal, 2004), or it simply may demand trust (Nonaka, 1994; Kogut and Zander, 1996; Hansen, 1999; Tsai, 2001). As innovation works on the technological frontier by definition, collaborating is particularly difficult since knowledge is neither codified nor understood by many. Organizations, thus, are the social locus where difficult-to-transmit knowledge can be exchanged, combined, and collective processed (Kogut and Zander, 1992; Singh, 2005). Firms provide a sense of community that allow individuals identify with the organization, adopting conventions, rules, argot and a way of thinking that enhance the ability for coding and decoding knowledge. As result, organizations generate collective intellects by strongly connecting people's intelligence (Kogut and Zander, 1996).

The collaborative network that constitutes the firm largely exceeds the formal organizational structure (Galbraith, 1977; Guimera et al., 2006a; Hansen and Løvås, 2004; Stevenson, 1990; Tsai, 2001). The whole internal social system can be a very complex web of interactions where knowledge and information are continuously flowing through. The structure of this complex machinery is neither fully manageable nor perceivable, but it certainly constitutes a firm's intellectual capacity (Levinthal, 1997). As it can determine the speed of spreading, the type of knowledge transmittable, the possibility of combining ideas and of collectively solving difficult problems, the exposure to information or the energy spent in keeping ties active among others (Katzenbach, 1993; Uzzi and Spiro, 2005; Granovetter, 1973; Tsai, 2001; Fang et al., 2010; DeCanio et al., 2000), the structure should achieve a delicate balance between density and sparsity. For dealing with high complex technologies, the collaboration network should be hierarchically near decomposable.

Hypothesis 7. *A firm's intellectual capacity will be positively related to the degree of hierarchical decomposability of its collaborative network of inventors.*

Since the complexity of systems is defined according to the size and level of interaction, it is expected that the degree of decomposability of a firm's social network plays a major role in explaining the firm's intellectual performance as its number of people is larger. Consequently:

Hypothesis 8. *The relation between the intellectual capacity of a firm and the degree of hierarchical decomposability of its collaborative network will be stronger as the collaboration network's size is larger.*

Finally, thinking an organization as a social cognitive machine whose information-processing mechanism can be described by the interaction between its cognitive map and the social network of its members, it would be expected that these two dimension complement each other. Both the collaboration network and the knowledge base can be viewed as two sides of the same phenomenon. While inventors are the physical components of the organizational machinery, the structure of the knowledge base would describe its abstract counter-part: the map of cognitive associations among theoretical concepts. Some authors have emphasized the right balance of the social structure of the firm in order to maximize diversity of knowledge and speed of transmission (Fang et al., 2010), others on the proper balance of the knowledge base (Yayavaram and Ahuja, 2008). However, as both of them are interrelated, their level of decomposability is expected to affect each other in a positive feedback. Without intending to introduce dynamic effects, it is simply expected a positive relation between the intellectual capacity of a firm and the interaction of decomposability levels of both the knowledge and the social structure. As Fang et al. (2010) states: "organizational intelligence can result from the accumulated wisdom of its members as well as from interactions among its members". While the first account for the knowledge base, the second accounts for the collaborative network. They are not independent dimensions but they are the very same machinery.

Hypothesis 9. *The interaction between the degree of decomposability of a firm's social structure and the degree of decomposability of its knowledge base will be positively related to the firm's intellectual capacity.*

As mentioned before, the empirical analysis will be performed with two different approaches. While the first introduces a explicit measurement of intellectual capacity of organizations, the second drops it for directly relating technological complexity, success and decomposability. This second analysis not only serves as a robustness control, but also allows testing another hypothesis.

When innovations are composed by very complex technologies, they are expected to fail in general with few exceptions that greatly succeed (Fleming and Sorenson, 2001). These expectations are based on a *general* behavior that does not consider the characteristics of the innovators involved. When considering the intellectual capacity of firms, this relation is not necessarily expected any more. Indeed, it would be expected to reverse: the more complex the involved technologies, the higher the expected success *if* the firm is capable enough. Coherently with Kauffman's model, as the interdependence of components increases, high peaks become fewer but higher. Therefore, a highly intelligent organization would be able to find those few highly successful configurations.

Specifically, if the firm is capable of decomposing the complexity of the technological landscape in simple patterns, the relation between interdependence and success of inventions should be positive; while if the firm is not, the relation reverses.

Hypothesis 10. *For low levels of decomposability of a firm's knowledge base, the relation between an invention's success and the level of interdependence of its components will be negative, while for high levels of decomposability will be positive.*

Hypothesis 10 works at a patent level explaining its success conditioned to the level of intellectual difficulty it poses to the firm according to the characteristics of the knowledge base and collaborative network of the organization. The statistical is approached in order to account for the level of interdependence of the patent.

In the following sections I explain in detail the chosen theoretical framework for studying the complexity of knowledge as well as the decomposability of both the social and knowledge structure. These concepts are calculated and measured for many firms and years within the Semi-conductor industry, which is well known in the literature because of its intense innovation, the high complexity of knowledge involved in their inventions, and the importance of patenting. Afterwards, it is statistically assessed the relation among them. A final appendix expands some aspect of measurements and statistical results.

4.3 Empirical setting

This paper focuses on the innovative dimension, specifically those inventions that are patented. Patents are legal mean for intellectual protection which indicates the potential commercial value of those innovations registered in it. Consequently, they contain very detailed information about the characteristics of the innovation, including inventors, firms, types of technologies involved, dates, and citations. To maximize the use of this valuable information, this paper analyzes firms operating in the Semi-conductor industry. In this sector, both the innovative activity is essential and the majority of innovations are patented (Hall and Ziedonis, 2001). It is well known in the literature that in this sector firms thoroughly endeavor to innovation as an essential activity for succeeding. That is why, innovating requires the maximum of their ability to explore and process knowledge. Human capital and research and development budgets (R&D) are not enough then. It is the whole network of inventors and their collective knowledge that is challenged at their maximum capacity when innovating. This paper tries to capture the capacity of organizations of innovating high with complex technologies, as well as how their social and knowledge structures, in their capacity for decomposing the complexity, are related to it.

The list of organizations that operate in the semi-conductor industry were selected from those enlisted in the Compustat Database (firms traded in the U.S. stock market). Out of active and non-longer existent companies, 595 were selected by the the SIC industry code 367. This code includes Electronic Components & Accessories (3670); Printed Circuit Boards (3672); Semiconductors & Related Devices (3674); Electronic Coils, Transformers & Other Inductors (3677); Electronic Connectors (3678); and Electronic Components (3679). In order to analyze their innovative activity, all firms operating in the semi-conductor industry were identified along with their produced patents out of The NBER U.S. Patent Citations Data File. For getting more fine-grained information about technological classes and sub-classes, complementary information about patents was obtained from the Harvard Dataverse Network's *Patent Network Database*. Finally, this information was checked, corrected and completed by a systematic direct consultation by algorithms to the official website of the USPTO for each patent in the database.

Class	Relative %	Cumulative %	Name
438	16.10	16.10	SEMICONDUCTOR DEVICE MANUFACTURING: PROCESS
257	9.58	25.68	ACTIVE SOLID-STATE DEVICES
365	6.38	32.06	STATIC INFORMATION STORAGE AND RETRIEVAL
327	4.26	36.32	MISCELLANEOUS ACTIVE ELECTRICAL NONLINEAR DEVICES, CIRCUITS, AND SYSTEMS
439	3.18	39.50	ELECTRICAL CONNECTORS
326	2.67	42.17	ELECTRONIC DIGITAL LOGIC CIRCUITRY
711	2.46	44.63	ELECTRICAL COMPUTERS AND DIGITAL PROCESSING SYSTEMS: MEMORY
375	2.24	46.87	PULSE OR DIGITAL COMMUNICATIONS
710	2.24	49.11	ELECTRICAL COMPUTERS AND DIGITAL DATA PROCESSING SYSTEMS: INPUT/OUTPUT

Figure 4.2: Most populated main technological classes in the semiconductor industry.

Following the general approach of Yayavaram and Ahuja (2008), I selected the 100 (out of a total of 337) technological classes where firms patent the most which represent 96.3% of the total data set. As Table 4.2 shows, there is high concentration on those technological classes directed related to the semi-conductor industry: only 9 technological classes accumulate 49.11% of the total sample of patents. Although this concentration, I take 100 classes so I can include wider interactions between more diverse knowledge components.

The USPTO divides technological major classes into sub-classes which are also hierarchically structured into 9 levels. In order to expand and capture more variance of technological characteristics, I consider three levels of classification: classes, level-1 sub-classes ("aggregated sub-classes" from now on), and sub-classes at maximum detailed level (simply "sub-classes" from now on). Finally, the database contains, 114,834 patents classified in more than 420,000 sub-classes, 149 firms in 27 years, information from 1976 to 2010, 100 technological major classes, 1,836 aggregated sub-classes and 22,344 sub-classes.

4.4 Dependent variable

This paper will address the analysis of the intellectual capacity of firms for navigating the technological landscape with two different approaches. First, it will be proposed a novel and explicit measurement of the intellectual capacity as dependent variable of a regression model that will use as main explanatory variable the decomposability of the firm in both its social and knowledge structure. The second approach will use patents as observations (not firms) and as dependent variable, the number of citations.

The construction of the dependent variable on the first approach, the intellectual capacity, demands careful explanation as it involves the processing of a large amount of data. Next sections describe how this measurement is calculated. Section 4.4.1 explains why the level of interdependence of innovations is considered as a measurement of intellectual difficulty and how it is calculated. Then, Section 4.4.2 constructs the performance space, a 2-dimensional space where patents are located according to their level of interdependence and their level of success. Since there are many sources of variation of patents' success (in term of citations), this Section also explains how success is considered as a standardized deviation from the expected mean of citations. Finally, Section 4.4.3 proposes the measurement of the intellectual capacity.

4.4.1 First step: Complexity of technologies

The first concept that must be defined is the *complexity* of knowledge. Many competing approaches and frameworks can be called for describing such an ambiguous concept. In this paper, I base the analysis on the adaptation of Kauffman (1993)'s *NK* model to the technological field by Fleming and Sorenson (2001). Understanding innovation as the result of combining prior knowledge, they state a parallelism between an organism's genome and an invention's patent. While genomes are composed by minor units of information called genes, patents can be seen as composed by different pieces of knowledge represented by the USPTO's technological classes.

Originally introduced for modeling the evolution of genomes, the NK model was proposed for analyzing systems made of several interacting components, where each component can be in one of many possible states. Genomes are composed by a complex structure of genes. These not only affect the design of the organism by its presence, but also in their relation with other genes. Some genes can suppress or alter the functions of others, and in that way, the interaction among them will ultimately affect the design of the organism and its ability to survive. By considering the number of components and their interdependence, it is possible to describe ex-ante the difficulty of the organism to mutate towards better fitted designs.

Similarly, innovations can be seen as genomes that combine N pieces of knowledge with some K interdependence. As N increases, the impact on the success of changing one component is smaller. On the other hand, as the interdependence of components increases, the variations on the performance will increase as we change singular pieces. This idea can be represented by Kauffman's model. Different innovations can be located on this theoretical landscape where the altitude of the surface indicates success (fitness in biological terms) and the distance among two innovations (locations) represents their dissimilarity.

The topography of the landscape, then, describes the difficulty of navigating those combinations in the quest for the highest peaks (the most successful innovations), and it depends on both N and K (Figure 4.1). While the first mainly determines the number of possible locations in the landscape, the second determines the roughness of the geography. A low level of interdependence describe a smooth landscape where mutations lead to marginal changes on the fitness level, while high level of interdependence implies abrupt fitness variations for any mutation on the design.[3] As the landscape become rougher, for a shortsighted navigator, finding the highest peaks using local clues will be more difficult. This is the key idea I use for capturing the intellectual difficulty of innovating. The complexity of technology, then, is understood as the difficulty of finding successful combinations when the components have a high level of interdependence.

[3]Formally, low levels of K describe landscapes with few and distant located maxima with a highly correlated spacial structure while high levels of K indicate many and much closer maxima with not spatial correlation. Furthermore, this landscape has an average fitness value inferior than for low levels of K, but with maxima higher that the maxima for low K's.

In their seminal paper, Fleming and Sorenson (2001) stated that *if* the NK model applied to the technological landscape, it should be observed an inverted-U relation between the level of interdependence and the level of success of patent. Their empirical analysis supported the existence of a maximum expected success at intermediate levels of interdependence among components of an innovation. However, the nature of the expected failure at low levels of interdependence is not the same that the one at high levels. While the first is a technical impossibility since there are not highest peaks to find, at high levels of interdependence the expected failure is due to the complexity of the landscape. In this scenario there are higher peaks that at intermediate levels of interdependence but those are more difficult to find. That is why Fleming and Sorenson (2001) also proposed and empirically corroborated that while the expected success of inventions decreases at high levels of interdependence, the variance of success increases (while it decreases at low levels of interdependence). Summing up, while the nature of failing at low levels of K is the strict absence of better configurations for an innovation, for high levels of K the nature of failing is based on the lack of capacity of dealing with complexity. At these scenarios there is plenty of scope for valuable innovations but they are extremely difficult to find since any minor change in their design leads to high variations of success. This is the reason why the intellectual capacity of organizations is evaluated at high levels of interdependence rather than at low levels.

For applying (Kauffman, 1993)'s theoretical framework, I follow the methodology of Fleming and Sorenson (2001). Patented inventions are considered as combinations of technological classes (types of knowledge). If patent p is classified into classes c_1 and c_2, it means the inventor(s) of the patent not only mastered those categories of knowledge but they were also able to combine them into a functional device approved by the USPTO (otherwise is not registered in the dataset).

Then, for defining the interdependence of an invention, first, it must be defined the *ease of combination* of technological classes. This idea is captured by the co-classification. Classes that are usually assigned to patents along with many others classes are assumed to be easily combinable. Then, the ease of combination is calculated as the average number of classes a focal class is assigned with. By considering the entire universe of USPTO's patent data set, for class c on year y is calculated as the mean of the number of classes class c has being classified with, in all patents that contain it with application date before year y. Equation 4.1 shows the latter:

$$E_{c,y} = \frac{\#\text{classes combined with class } c}{\#\text{patents classified in class } c} \tag{4.1}$$

If patents assigned to technological class c are also classified along with many other classes, then, c is assumed to be easy to combine and therefore it does not interact much with the other categories. A patent classified in c, thus, would be a combination of technologies located at a smooth landscape (if the other classes have a similar degree of easiness of combination).

For each of the 22,344 sub-classes in the data set from 1980 to 2003, the ease of combination is calculated. It was used the entire universe of patents contained in the Harvard Dataverse Network's Patent Network Database discarding design (D), statutory invention registration (H), plant (P) or reissue patents (R); a total of 3,984,771 patents classified in 17,445,405 classes/subclasses.

Since technology mutates over time along with the use of technological classifications, the easiness of combination is calculated yearly. For each subclass and each year, I consider all patents classified in the focal sub-class with application date earlier or contemporary to the focal year. The application date was considered instead of the granting date since the first is closer to the moment the combination of knowledge was done and the invention was finished and considered worthy of being registered. On the other hand, the granting date informs the moment a patent is approved by the USPTO and this date can be substantially different from the previous one. Considering application dates does not mean considering applications. All patents considered in this paper are all granted patents.

The ease of combination is used to calculate the interdependence of a technological class. Kauffman models the interdependence as the K parameter which indicates how much elements interfere between them in increasing or decreasing their individual contribution to the organism's fitness. Specifically, K states how many elements of the string are affected by each of them. Considering technological classes as the pieces that conform an invention, then, if a patent p is classified into N classes, it is averaged the easiness of combination of each class and inverted to calculate the interdependence coefficient of this particular patent (equation 4.2).

$$K_p = \left(\frac{\sum_{c=1}^{N} E_{c,y}}{N} \right)^{-1} \tag{4.2}$$

As the responsible of describing the difficulty to find high peaks in the technological landscape for a constant N, the K-parameter can be interpreted as an independent measurement of the intellectual challenge that a particular patented invention represents.

4.4.2 Second step: The performance space

The level of interdependence among innovations' components describes the difficulty for perfecting as well as it indicates the possibility of some very useful configurations (Fleming and Sorenson, 2001). As such, it will be used for measuring the intellectual challenge of innovating. The roughness of the technological landscape illustrates the difficulty of navigating it for looking optima, specially with the lack of spacial clues. On smooth landscapes, as the one in the left on Figure 4.1, following the slope is enough to find the absolute optimum. On the contrary, this strategy only leads to a local optimum on a roughed scenario (right side of Figure 4.1). However, different navigators could have different horizons of perceptions. Therefore, while a shortsighted innovator could end up trapped into a local optimum, other might be able to see further and jump to other peaks. In other words, the complexity of landscape determines the potentiality of success but it depends on the ability of the innovator to succeed on it. If all firms had the exact same capacity for innovating, it would be expected a decrease in success on average as K increases. However, this does not necessarily holds as we contemplate different intellectual capacities for navigating the technological complexity.

If we combine the information about the level interdependence of an innovation and its success, it is possible to observe, by comparison, the different intellectual capacities of innovators for dealing with complexity. In this two-dimensional space of technological complexity and success, the intellectual capacity of organizations can be captured. While the measurement of a patent's interdependence was explained in the previous section, the commercial success of inventions will be measured as a standardize and controlled count of citations.

For each patent, the number of citations received during the 5-years period after its granting date is considered as the measurement of technological success of the invention. Although the innovation literature recognizes a correlation between patents' citation rates and their commercial success (Hall et al., 2001), I explicitly state *technological* success rather than other. This is mainly because I analyze inventions in a biological framework where, for the sake of keeping the parallelism, successful designs are those which survive and leave their legacy in their offspring. Highly cited patents success in terms of their "offspring", i.e. inventions that use the knowledge embedded in it.

In order to collect data about citations, an algorithm sequentially extracted the information about citing patents and their dates of application directly from the US Patent & Trademark Office, Patent Full Text and Image Database online. I consider those citations that take place within the 5 years after a the *cited* patent is granted by considering the application date of the *citing* patents [4]. Given the length of the considered time-windows, I discarded those patents that are not old enough to account for this period of citation in the data set (the last date of citations is 2011 October).

The basic idea is to standardize citations but considering not only its observed mean and variance, but considering a comprehensive set of statistical controls. Since citation is a count variable (discrete and non-negative), a negative binomial regression with modeled variance is used to standardize its values by including the following controls.

Controls for standardizing cites

First of all, citations should be controlled by the *number of components* inventions have. Circumscribed in Kauffman's model, both interdependence and number of components determine ex-ante the nature of the design landscape. When facing the same interdependence, an invention with fewer pieces to combine represents a greater challenge than one with a large numbers. Thus the number of components must be taken into account when comparing inventions' success.

A special case calls for a different treatment: innovation with only component. A dummy variable is included to distinguish those inventions since they do not combine different technological sub-classes and the complexity of technology is defined upon the difficulty of combining.

Besides the N parameter from Kauffman's model, other controls should be included in order to differentiate characteristics of the inventions that might influence the citation counting after their publication. First, I include the number of *prior citation* as a measure of the localness of search. By simply counting the number of references to other patents, this variable controls for the scope of the local search and the propensity of the sector for citing (Fleming and Sorenson, 2001). Furthermore, it also captures the combinatorial problem of an invention but from another perspective. Instead of combining technological classes, patents combine knowledge in other patents.

Another control, the number of *previous trials*, accounts for the number of previous combinations of the exactly same set of technological sub-classes the focal patent combines. In other words, it counts how many inventions has been filed before combining

[4] For capturing this period of citations after the patent is granted, the minimum unit was the month. The algorithm was run on August 2013 (the on-line USPTO data base may have modified since then)

the same technologies and therefore, it controls for a possible declination in technological success of invention that are basically replica of components and structure. Furthermore, since each patent represents a local peak in the technological landscape, this variable also complements the parameter K in describing the local roughness (Fleming and Sorenson, 2001). For each patent and the sub-classes it couples, the variable counts how many patents in the USPTO has been filed before the focal patent and that are also classified in the same technological sub-classes.

In a similar line of reasoning, two other controls are considered to capture previous trials but this time with regards to the history of the firm. These two would be closely related to *search depth* (Katila and Ahuja, 2002), i.e. the degree of reuse of firms' knowledge.

When combining different technologies to innovate, the firm must explore different ways. The success of a patent, thus, would be affected by its previous attempts. As the firm learns and becomes familiar with certain technologies, the navigation of the design landscape should be easier. That is why I control, first, for *sub-classes already used by the firm* by counting, for each patent, how many of the technological sub-classes used in a patent are already used in the firm's knowledge base. Since the number of sub-classes of patents varies, the number of already used sub-classes is divided by the total one. Therefore, this measurement lies within the [0-1] interval; 0 meaning that the firm has not used the technological sub-classes during the previous 3 years, and 1 that the firm has used at least once all the sub-classes the focal patent is classified in.

The second control regarding the previous experience of the firm regards the *couplings already used by the firm*. The firm may have already not only used the sub-classes involved in a patent, but also it may have combined them. Given that a patent is a combination of technological sub-classes, I control for previous combinations by counting how many couplings of those sub-classes has been done in the firm's knowledge base out of the total of coupling the focal patent has. For instance, if a patent couples 4 sub-classes - A B, C and D -, then, there are a total of 6 couplings involved. If the knowledge base of the firm reveals to have coupled only sub-classes A and B, then, the value of this variable is 1/6. Logically, this variable also lies within the [0-1] interval.

The number of inventors involved in the patent may also affect the difficulty of navigating certain technological landscape. Two inventions with identical interdependence may differ in the number of people directly assigned to the project and this affects the propensity to be referenced by upcoming inventions. Controlling for the *number of authors* registered in a patent is extremely pertinent when the theoretical base of the entire paper claims that organizations, as social arrangements, arise in order to increase the capacity to deal with high complex knowledge. Consequently, this variable is included by accounting the number of inventors that formally worked together developing the focal patent.

While the set of previous controls have a theoretical reason to be considered in Kauffman's framework, the following controls are included to isolate citations for other sources of variations. In the first place, the number of citations might be influenced by the nature of the technologies and inventors operating in different industries and sectors. *Technological main class*, thus, is introduced as a control. This paper considers the top most populated 100 USPTO's technological classes where firms in the semi-conductor industry work in. Since classes cover highly differentiated technological areas, citations rates across differ enough so as to control for this source of variation. Legal strategies, the technical nature of the technology, the propensity to patent on the sector, and complexity and scale of the technology are some of the reasons for adding this control.

Even though controlling for main technological class, the number of main classes is also considered. Being classified in more classes certainly affects the probability of being cited since it exposes the patent to different technological areas. Furthermore, the number of classes also indicates the scope of the invention. I include the *number of major classes* for capturing the level of exposure and therefore, the changes in the propensity to be cited.

Since rates of citation may also change across time, I control for *granting year*. Even though for all patents is considered a period of citations of 5 years after the patent is published, the granting date must be taken into account. The main reason lies on the lag between the filing and granting date. When counting citations, only those patents with filing date within the 5 consecutive years after the publication date of the focal patent are accounted since I consider that patents are generated at the moment of the filing date. Because of this, citations of newer patents might be biased: there are filed patents that cite the focal one but since those are not yet published I cannot know whether these citation lies on the citation window. Furthermore, citation might be influenced by other factors that change over the years and therefore they are controlled in this variable.

After including all these controls that may impact on the citation rate of patents, Table 4.1 shows the coefficient of the generalized negative binomial regression model while Table B.6 in the Appendix shows descriptive statistics of the involved variables.

Table 4.1: Estimated coefficients for controlling citations

Generalized negative binomial regression
Dependent variable: citations (5-years window)

Dependent Variable	**mean**	**variance**
constant	0.6395 (0.107) ***	0.1306 (0.199)
prior citations	0.0095 (0) ***	0.0006 (0) **
number of authors	0.063 (0.003) ***	-0.0116 (0.003) **
number of major classes	0.0575 (0.007) ***	0.0028 (0.008)
previous trials	-0.0003 (0) +	-0.0007 (0) **
couplings already used by the firm	-0.014 (0.012)	-0.0272 (0.016) +
sub-classes already used by the firm	0.0752 (0.013) ***	-0.0906 (0.017) ***
number of sub-classes (N)	0.0331 (0.002) ***	-0.0176 (0.003) ***
single sub-class	-0.0448 (0.016) **	0.1224 (0.021) ***
dummies main class included		
dummies granting year included		
	*** $p < 0.001$ ** $p < 0.01$ *$p < 0.05$ +$p < 0.1$	
Num obs	110,693	
Prob chi2	0.0000	
Log pseudolikelihood	-361,264.92	

The estimation of this model aims to isolate the variation of this citation from other sources than complexity. By estimating the mean and variance for each observation with this model, the number of citations is standardized as a deviation from the mean. In order to avoid abnormally deviated observations, they are winsorized at 99% of its range of variation. The resulting range of variation is then scaled into the unity. Finally, the success of each patent is observed as standardized deviation from the mean explained by all the controls but the complexity. I call *intellectual achievement (IA)* to the remaining unexplained variation of a patent's success.

4.4.3 Dependent variable: Intellectual capacity

The saying is that brilliantness and madness only differ in the degree of success. If you are able to solve an extremely difficult problem, you will be called a genius while if you fail you will be called crazy since you thought you were capable of overcoming such challenge. The psychometric perspective of intelligence considers it as the capacity for thinking and solving a problem (Spearman, 1927; Glynn, 1996) and this is what I try to replicate on this Section.

When innovating an organization deploys its intelligence for comprehending the complexity of technology. This capacity cannot be tested in a controlled experiment but

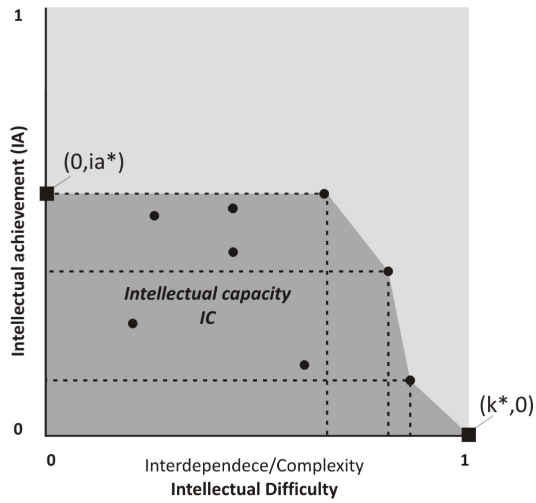


Figure 4.3: Difficulty-success space

it can be observed since innovations are natural tests. For assessing an individual's intelligence, s/he is usually exposed to a series of problems of progressive and controlled difficulty with solutions that can be either correct or incorrect. Mimicking this logic, I consider innovations as intellectual problems firm deal with, where interdependence of their components measure the difficulty, and their level of success evaluates the "correctness" of that solution (what I have called before *IA*).

Being *IA* one of the dimension of the performance space where firms showed its ability to deal with complex knowledge, the other one is the K parameter of Kauffman's model. As previously explained, this captures the nature of the landscape regarding the height of peaks, the amount of them, and the spatial correlation that make them easier to find. The higher the interdependence among components, the most difficult to perfect innovations (Fleming and Sorenson, 2001; Ethiraj and Levinthal, 2004)

The 99% of variation of the interdependence parameter K is also circumscribed to the unity after winsorizing. The resulting space is contained in the unit square and it provides the difficulty-success space needed to capture the capacity of firms to innovate with complex technologies. Every year, firms find several combinations of technologies that are patented with different levels of success, as Figure 4.3 shows. As an IQ test checks what is the maximum difficulty a person can solve, I will look for those patents of certain firm and year that are not dominated by any other in terms of difficulty and success simultaneously.

Given that *IA* is not binary but a continuum, I look for the highest peaks a firm can find at different levels of technological roughness. Differently to a intelligence test, the firm may produce several innovations with the same level of intellectual difficulty during the same year. Therefore, only the *most successful* innovations are considered to evaluate the intellectual capacity of organizations as if they had several trials to show what they are capable of.

Considering that a patent dominates another one whenever it has achieved a greater success at a rougher technological landscape, I look for those patents that are not dominated by any other. Based on this frontier, I measure the ability of a firm to innovate in high complex scenarios as the surface below it (dark-gray area on Figure 4.3). For that purpose I also assume that generating a patent at the maximum level of difficulty has zero success and at zero interdependence has the maximum success achieved by the firm (black-squares at Figure 4.3).

Defining $p_i = (K_i, IA_i)$ as a patent located in the performance space, $P_{f,y}$ as the set of all patents produced by firm f during year y, and $k^*_{f,y}$ and $ia^*_{f,y}$ (Figure 4.3) as:

$$\begin{aligned} k^*_{f,y} &= max(K_p)/p \in P_{f,y} \\ ia^*_{f,y} &= max(IA_p)/p \in P_{f,y} \end{aligned} \tag{4.3}$$

Then, I define:

$$P^*_{f,y} = P_{f,y} \cup \{(k^*_{f,y}, 0), (0, ia^*_{f,y})\} \tag{4.4}$$

Finally, the intellectual capacity of a firm f for year y is defined as:

$$IC_{f,y} = \text{convex hull of } P^*_{f,y} \tag{4.5}$$

Summing up, taking one year of a firm's patent production, considering the level of interdependence of those patents and their success after controlling for a set of measurements, it is calculated the intellectual capacity of a firm on that year as the surface below the most successful and most complex patents it has produced. As the variation of both *IA* and K of a patent have been circumscribed to the unity, the range of variation of the intellectual capacity of firm manifested in its innovations also ranges between 0 and 1.

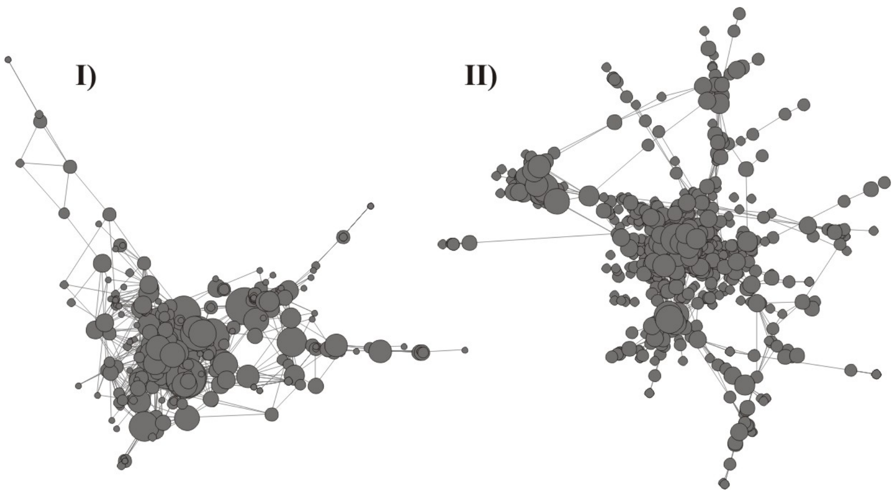


Figure 4.4: Intel's technological sub-classes coupling network - knowledge base (left panel) and co-invention network (right panel). The size of the nodes is proportional to the number of patents assigned to it.

4.5 Independent variables

Once defined the intellectual capacity of an organization as an observed performance when creating new technologies, I relate it with its ability of decomposing complexity as manifested in its knowledge base and its collaborative network. Using information registered in patents, both the knowledge base and the collaborative network are constructed for each firm each year on the Semi-conductor industry.

Following Yayavaram and Ahuja (2008), the knowledge base is defined as the knowledge a firm manifests to have managed during the three years previous to the measurement of performance. Those patents generated by the firm during this period (according to the application date) are considered along with their technological classification. Considering technological aggregated sub-classes of aggregation according to the USPTO, I construct the coupling network. A patent is considered to couple a pair of sub-classes whenever it is classified in both of them.

The knowledge base constructed this way shows not only which types of technologies a firm masters but also how they are related by the firm. Frequent patenting into some technological class indicates a experience on this field. The frequent coupling of a pair of classes, on the other hand, suggest the firm understands those two as a compatible and promising combination of knowledge (Yayavaram and Ahuja, 2008; Katila and Ahuja, 2002). Furthermore, the coupling structure can be assembled into a network that also reveals indirect relations among technological classes as the firm understands them.

On the other hand, the social structure of firm's inventors is proxied by the co-authorship network. Looking into the same patents selected for the knowledge base, instead of considering their technological sub-classes, I analyze their authors. Those inventors that are registered in USPTO file as co-author of a patent are considered to have a collaboration tie.

The co-authorship network provides some rudimentary but valuable information about the structure of collaboration among inventors working for a firm. The literature on innovation usually considers a co-authorship tie as social relation where a considerable amount of knowledge flows trough given that inventors have worked together to develop the invention (Phelps et al., 2012). The strength of the tie, then, is suggested by the number of patents inventors have co-authored. The more they collaborate, the stronger the channel. When all ties are simultaneously considered, the entire collaboration structure of the firm is revealed and it can depict interesting features of how knowledge is collectively processed within the organization.

As an illustration of these two concepts, Figure 4.4 shows the knowledge structure of Intel from 1992 to 1994 on the left panel and the co-authorship structure for the same period on the right panel. On the knowledge structure nodes represent technological sub-classes whose size is proportional to number of patents classified on them. The intensity of couplings determines the proximity of nodes using a force-directed graph. The nodes of the social network, on the other side, represent inventors whose size is proportional to the number of patents they have authored. As in the knowledge network, the strength of co-authorship ties determines the proximity of nodes using the same algorithm for plotting.

These two networks provide information about how organizations structure both their cognitive map and intellectual collaboration. Both of them can be represented by a square and symmetrical matrix, where each row/column depicts the connections of a technological sub-class in the knowledge base or an inventor in the collaborative network. As such, they depict the way firm decompose the complexity of technology into interconnected clusters of knowledge and inventors. Simplifying complex phenomena into a hierarchical nearly-decomposable structure mimics a cognitive strategy of human minds. That is why, based on the Simon's approach, I propose a way of measuring how much decomposable are these two structures.

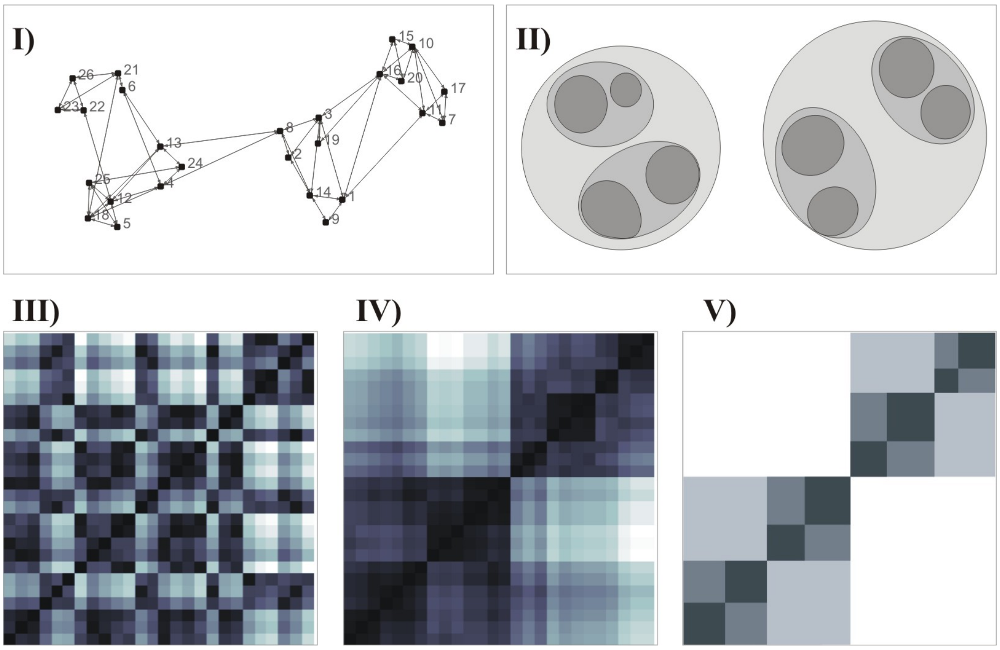


Figure 4.5: Illustration of a network hierarchically clustered in 3 layers. Panel I depicts a network of 26 nodes where the length of ties represent the intensity of connections. Panel II shows how nodes group in three layers according to the intensity of relations. Panel III shows the coupling matrix of the network where the darkness is proportional to the intensity of ties. Panel IV shows the same matrix sorted. Panel V shows the simplification of the matrix into hierarchical clusters.

4.5.1 Knowledge base's hierarchical decomposability

Simon defined hierarchical systems as those composed by interrelated sub-systems, each of the lagger being, in turn, hierarchical in structure until it reaches some lowest level of elementary subsystems (Simon and Ando, 1961; Simon, 2002, 1962). Specifically, he defines a hierarchic system according to the patterns of zeros and near-zeros in a matrix P that has the following structure:

$$P = \begin{pmatrix} P_1 & & & & \\ & \ddots & & & \\ & & P_i & & \\ & & & \ddots & \\ & & & & P_n \end{pmatrix} + \varepsilon C \tag{4.6}$$

While P_i's are square sub-matrices and the remaining elements are all zeros, C is an arbitrary matrix of the same dimension and ε is a very small real number. P, then, is a nearly decomposable matrix. The intuition behind this structure is the distinction between the interaction $among$ subsystems and $within$ subsystems. The interaction among subsystems, even though positive, is expected to be considerably weaker than the interaction within subsystems. The matrix P should show very low values outside the blocks of the main diagonal ($P_i's$).

When a system has a hierarchical architecture with more than one level of clusterization, the matrix of interaction should show a structure like the following one (Simon, 2002):

$$P = \begin{pmatrix} & \mathbf{p} & \varepsilon & \varepsilon & \eta & \eta & \eta & \eta \\ \mathbf{p} & & \varepsilon & \varepsilon & \eta & \eta & \eta & \eta \\ \varepsilon & \varepsilon & & \mathbf{p} & \eta & \eta & \eta & \eta \\ \varepsilon & \varepsilon & \mathbf{p} & & \eta & \eta & \eta & \eta \\ \eta & \eta & \eta & \eta & & \mathbf{p} & \varepsilon & \varepsilon \\ \eta & \eta & \eta & \eta & \mathbf{p} & & \varepsilon & \varepsilon \\ \eta & \eta & \eta & \eta & \varepsilon & \varepsilon & & \mathbf{p} \\ \eta & \eta & \eta & \eta & \varepsilon & \varepsilon & \mathbf{p} & \end{pmatrix} \tag{4.7}$$

Where $\mathbf{p} > \varepsilon > \eta$. P, then, is formed by 4 blocks of intensity $\mathbf{p}$ that are grouped in a higher level in two blocks with a level of interaction among these blocks of η. The system represented by P is a near decomposable system of 3 layers.

Even though the intuition of this idea is simple, Simon never defines a way of measuring $how\ much$ hierarchically decomposable a system is. Yayavaram and Ahuja (2008) proposed a modified network clusterization measurement for capturing the near-decomposability of firm's knowledge base. However, what they measure is how much networks cluster in *one* layer. This operationalization of the near-decomposability, even though it is very valuable as a descriptive measurement, it does not fully incorporate the overall concept: the hierarchy. Simon proposed that highly complex systems can be much easier to understand and to analyze whenever they show a hierarchical near-decomposable structure. The idea of many layers of subsystems within subsystems that are also nearly decomposable is the one that explains all the advantages of this structure.

Here I propose a measurement for capturing *hierarchical decomposability (HD* as theoretically proposed by Simon that allows comparing different systems. Basically, this measurement attempts to assess how *simple* is a knowledge base's structure. Its size and interactions can be ordered in a HD structure that allows focusing the limited cognitive capacity in some parts of the system without loosing track on the rest. In order to avoid overloading this section, the methodology is sketched in the following paragraphs but it is explained with full detail in Section B.1 of the Appendix.

In order to capture the "simplicity" of the system, this measurement asks how difficult finding a component would be when the location in the structure is unknown. If the system was zero decomposable, there would not be any distinction or pattern among relations between elements and the system should be analyzed in its totality. However, if the system presented some organization, it would be simpler and therefore it would require less analysis. Thinking the knowledge base as a library, the number of sections, its hierarchy and the ambiguity of the division would determine how difficult finding the best book we need is (which we do not know exactly but we must study its content). We can analyze sections as unities, quickly discarding entire sets of books, and once identified the group where the element might be, it is only this section the one which is analyzed holistically. As we add layers of sub-groups, it is easier to quickly find the wanted element.

Three variables enter this analysis. First, the number of layers or levels the system has; second, the number of groups that form each layer; and third, the level of decomposability of each layer. For analyzing these features, the measurement is based on the dissimilarity matrix among the technological sub-classes from a firm's knowledge base. Instead of analyzing sub-classes as connected or disconnected, this approach considers them as located in a relational space where closely located classes are highly coupled by the firm and the other way around. To define distances among couples of classes in the network there is not need of those to be directly connected but they can be defined by considering indirect connections. For instance, if class A is usually combined with class B when generating patents, and class A is also usually combined with class C, while C and B are not together in any patent, still, C and B will be closely located.

The knowledge structure can be represented by a symmetrical matrix whose $(i, j)^{(th)}$ element counts the number of patents that are classified simultaneously in classes i and j. These coupling intensities are interpreted as a measurement of proximity in a relational space. As such, they can be transformed into a measurement of distance or dissimilarity. By using all the information available in dissimilarity matrix obtained by the coupling network (knowledge base) geodesic distances are calculated among all connected (directly or indirectly) elements. If some elements of the structure are unconnected, the maximum distance in the entire set is assigned to those couples (Section

B.1). Afterwards, this matrix is agglomeratively clustered by minimizing the average distance within groups (equivalent of maximizing average intensity within groups) in order to reveal its hierarchical architecture (Section B.1). The classification tree, then, provides information about the number of levels and the number of groups at each level that draw the structure of the knowledge base. This information is used for measuring the level of decomposability of each group within the structure for finally combining all these measurement into a global one that describes the decomposability of the entire knowledge base.

For each group its decomposability is measured in terms of how different element are *within* its sub-groups than *among* its sub-groups. If relations among groups are null, then, the structure is full-decomposable. On the contrary, if the relations among groups are of the same intensity of those within groups, the structure is zero decomposable as there is not difference of clustering. As the ratio of this two kind of relations is larger, it is easier to recognize patterns and understand the phenomenon. The absecence of structure implies analyzing everything at the same time while certain degree of decomposability allows focusing on some parts without loosing the entire picture. Specifically, the *decomposability of a group* ranges from 0 to 1. It is 0 when the group has no structure at all and it increases to 1 as the group decompose into homogeneous sub-groups but highly heterogenous among them.

The decomposability of the group is used to calculate its *simplification* as a weighted average between two possible scenarios. Weighted with the decomposability measurement it is considered the number of sub-groups within the considered group. If the decomposability would be 1, then, the complexity of the group would be entirely simplified to the number of sub-groups. On the other hand, weighted with (1- the decomposability of the group), it is considered the total number of elements without any clusterization. Thus, if the group is zero decomposable there is not clear distinction between elements and therefore would not be any simplification. If decomposability lies between 0 and 1, then, the simplification of the group is a weighted average between the reduction to few groups and the absence of any clusterization.

The same logic is applied to all groups in the hierarchical structure. Therefore, the whole simplification measurement of the knowledge base (and the collaboration network) is calculated by a recursive algorithm that starts calculating the simplification index for those groups of first order (those which gather individual elements) and continues by calculating the simplification measurement for second order groups (those which gather groups of elements) until it arrives to the maximum order group.

The global measurement, *simplification of the knowledge base (E)*, also ranges from 0 to 1 and it can be interpreted as how much the whole system is simplified throughout its structure. If a system of n elements measures 0.5, it means that analyzing the entire set of elements only takes 50% of the effort as a consequence of its hierarchical decomposability. However, the final measurement has two modifications before introduced in the statistical analysis. First, it is optimized. When dividing the system into agglomerative clusters, it might be a point where the gains of dividing sub-groups into sub-sub-groups might imply a lost in simplification. That is why, it is calculated at different levels so as to choose the minimum simplification achieved by the structure. Secondly, the final measurement will be transformed into the *hierarchical decomposability (HD)* of the knowledge base as HD =(1 - E). This is done so it is easy to interpret the signs of the regression analysis. The decomposability index will increase as the hierarchical decomposability of the system is better outlined, achieving high values when the system is hierarchically nearly decomposable. This measurement is independent of size but its variation is more sensitive to larger groups since there is more space for simplification. It is explained with full detailed in the Appendix along with Figure B.2, Figure B.3 and Figure B.4 providing examples.

4.5.2 Social structure's decomposability

When tackling problems of high complexity, individuals must focus their intellectual efforts in areas they can embrace and then collaborate with others in a proper environment that facilitates a clear communication channel. Logically, as they get larger, organizations should increase its whole intellectual capacity more than proportionally since the room for connections among members grows faster. However, a full connected structure not only is impossible for large groups but it would be extremely inefficient in terms of cost of keeping relations and because of the overexposure to circulating information. Similarly to the idea of *wiring optimization* of neuronal networks structures (Chen et al., 2006), the optimal functioning should lie between the full connected network and the absence of any interaction. Once again, the complexity of the system of inventors within an organization should decompose into a heavily clustered groups that allow fast transmission and processing of knowledge, and weak but existent connection across groups that foster the circulation of disparate knowledge and ideas. This delicate equilibrium between depth and breadth empower innovation the most (Chen and Guan, 2010; Cohen and Levinthal, 1990; Fleming et al., 2006; Kleinberg, 2000; Powell et al., 1996; Uzzi and Spiro, 2005; Watts and Strogatz, 1998).

The interaction among people withing an organization largely exceeds its formal structure and communication channels (Tsai, 2001; Tortoriello and Krackhardt, 2010). Thus, mapping a social network requires not only being able of capturing large amounts of data that are not usually easy to observe and register, but also defining what kind of social relation are relevant for the analysis. Some social interactions merely rituals of politeness, while others are the responsible of joining or processing key knowledge for innovating. As previously explained, this paper focuses on co-authorship invention as formally registered at the USPTO.

Co-authorship ties describe close collaboration among inventors when innovating. They are strong ties that heavily exchange and process knowledge on daily bases. By considering only this kind, this approach rules out weak ties, those of occasional nature, unable to transmit tacit, uncoded or high complex knowledge, but capable of joining distant parts of the social map and transmitting some information that might be critical to innovation (Granovetter, 1973).

It can be argued that in an industry such as the semi-conductor one, where the complexity of knowledge in all of the technology is so much, the role of weak ties might be minor. However, even though co-authorship ties can be seen as strong ties, not necessarily all strong ties are co-authorship ties. Therefore, the co-authorship network is expected to capture not all but exclusively strong ties within an organization. Needless to say, this approach also rules out those firm's members that are not inventors.

Although its limitations, this approach stills captures valuable information on the social structure of the firm as its wide used in the innovation literature proves it. By capturing the entire network of co-authorship it is possible to reduce the impact of missing information about collaborative ties. Second, the network of inventors represents a sub-set of firms specifically advocated to research and development. Third, the frequency of co-authorship differentiate among intensities of ties. Some inventors usually work together while other occasionally. Even though both relations are strong enough to support the activity required for an invention, their different intensities have an important impact on the overall network.

As developed before, the complex nature of the problems in the semi-conductor industry leaves not option but gathering and connecting many capable minds in order to deal with them. And as the size grows, the structure and pattern of connections matters the most. As previously referenced, organizations outperform markets in dealing with high complex knowledge as they provide their members the right social environment. However, social structures are not fully manageable and they might seriously affect the intellectual capacity of processing knowledge. Atomistic specialization coordinated by an organization is not enough. When the size of the company is large enough, the complexity of the network of interaction might not fully harvest the gains of being organized.

For analyzing this possibility, the co-authorship network is constructed based on the information provided by the USPTO. Directly downloaded from the on-line data base, the names of inventors of each patent are considered to build the co-authorship network contemporary to the knowledge base. Mapping this network usually presents the problem of matching names, a variable that does not posses enough variation and therefore can easily misidentified different elements as the same. As this paper analyses social networks in a 3-years windows of inventors that belong to the same firm, then, this possibility is negligible. The network, thus, is composed by inventors connected by the number of times they co-author a patent (Figure 4.4 provides an example on its right panel). Differently to the knowledge base, instead of coupling technological classes, this network couples inventors.

With this information, I proceed to calculate social distances among all the network's inventors. Distances, in this case, indicate the likelihood or easiness for knowledge to be transmitted among people (Singh, 2005). Once the matrix of distances is obtained, the whole set of inventors is hierarchically clustered for finally calculating the decomposability of the whole network as explained in Section 4.5.1 and expanded in the Appendix B.1. Similar to the knowledge base, the decomposability of a firm's network of inventors indicates its simplicity. The more hiearchically near-decomposable the structure is, the closer to 1 will be the decomposability measurement. As organizations gather inventors to increase their capacity for innovating, it will be so as the social structure allows an efficient assembling of all of them in terms of flow of information and deep processing.

4.5.3 Controls

Since firms differ among each other not only in their knowledge base and collaboration network, other characteristics that might influence the intellectual capacity of firms should be considered. In the first place, I will control for the resources firms directly assign to research and development (R&D). Absorptive capacity is usually related to R&D investments (Cohen and Levinthal, 1990; Tsai, 2001). Even though it does not fully explain the intellectual capacity of an organization, team, or even a single person, it certainly influences as it mainly provides physical means, like equipment and facilities. R&D expenditures are considered as a a share of the firm's total assets. Both measurements are obtained from the Compustat Database.

When analyzing firms as complex social machineries, its size must be considered since its determines their potential complexity. Even though I use the co-invention network for assessing the social structure, the number of inventors and their collaboration ties do not capture the full *size of the firm*. Inventors are a minor part of an organization and there are many other social interactions among other jobs and functions within that form the whole firm. The logarithm of the total number of employees is considered, thus, to control for different thresholds of complexity.

Some researchers also propose controlling for the *performance* of the firm since profitable ones may not pursue innovation as hard as a less successful one. This control is incorporated considering return on assets.

The knowledge base is assumed to manifests what firms know. Following the approach of Yayavaram and Ahuja (2008), knowledge bases were considered in three-years windows. However previous experience might also influence the success of inventions, not only because of the accumulated knowledge but also because of phenomena like reputation or visibility. For these reason, in the statistical analysis is considered the *number of years that the firm has been patenting*.

The number of inventors who work for a company must be also considered. A greater number of inventors describe the potentiality of the social network in terms of diversity and collaboration. However, the size of the inventor staff is relevant relative to the size of the knowledge base. Because of this, the analysis includes the *ratio between inventors and the number of classes* of the knowledge base with the purpose of capturing how much human power is devoted to the different areas the firm deal with.

Even though intellectual capacity of firms is captured according to only the most successful patents they generate along one year, the rest of patents should be considered as well. According to the methodology explained in Section 4.4.3, only the best patents of firms are selected to calculate the maximum it can perform. The rest of them would be considered as trials and therefore it should be taken into account since a producing many or few patents could affect this measurement in a non clear way. *Ceteris paribus*, many patents might indicate a larger effort and a possible declination in success but it can also be associated with higher citations rates because of cross-citation or just because of the number of attempts. Either way, the *total number of patents produced in the focal year* is considered for differentiating this cases.

As the knowledge base and collaboration network are proposed to be associated with the capacity of firms of processing high complex knowledge, specially in the simplicity of its structure measured by the hierarchical decomposability of these structures, the *size of knowledge base* must be taken into account. This variable simply counts the number of different technological sub-classes a knowledge base' patents have been classified in. It indicates the variety of technological areas the firm has managed during the 3-year period windows.

According to the methodology explained in Section 4.4.3, only the best patents of firms are selected to calculate their intellectual capacity. The rest of them are considered as trials and therefore they must be taken into account since the number of attempts can certainly affect the measurement. *Ceteris paribus*, many patents might indicate a larger effort and a possible declination in success, but it can also be associated with higher citations rates because of cross-citation or just because of the number of attempts. Either way, the total number of patents produced in the focal year is considered in logs and it is called *patents*.

Furthermore, other measurements of structural characteristics of the firm are considered. Since both the knowledge and social structures are represented by networks, for both of them the level of *clusterization*, the number of *unconnected blocks*, and the *density* are calculated for controlling different characteristics beyond its level of decomposability. For the network of inventors is also calculated the *mean intensity of ties*.

Finally, in order to control for the possibility that firms choose the level of complexity of the technology they combine for innovating, the mean complexity of patents considered for calculating the *IC* is included. When firms innovate, patents register the level of difficulty they are dealing with and the level of success of this trial. If firm randomly choose the level of difficulty, or if they uniformly try different levels of difficulty, then, the *IC* would unbiased capture their capacity. However, it may happen that firm choose the level of difficulty and therefore, different measurements of IC cannot be compared. Organizations dealing only with low levels of complexity would never reveal their capacity when tested at higher levels. Because of this reason, the mean level of complexity of patents considered for calculating *IC* is included as control. The control variable *K-level* is aimed to differentiate those firms that systematically choose low complex technologies.

4.6 Statistical Analysis and Results

After defining the intellectual capacity of firms, the postulated association with firms's decomposability of the complexity and a proper set of controls, a linear regression model is performed. Because the dependent variable is calculated yearly for many firms, the data is analyzed in a Panel. A fixed effects model is used as it captures unobserved characteristics within the firm which may impact on its ability to process knowledge. Statistically, the Hausman test (Hausman, 1978) strongly rejects a random-effects model.

The theoretical approach of this paper also supports this model. The goal of the statistical analysis is to test whether there is a significant association between the intellectual capacity of firms and their level of decomposability in both its knowledge and social structure. Changes in the level of decomposability should be observed along with changes in the same direction of intellectual capacity of firms. However, the underlying idea of this approach is that organizations are differently capable of innovating with complex technologies and this is controlled by the Panel. By assuming the correlation between firms' error term and the measurement of intellectual capacity, a fixed-effect model removes the effect of those invariant characteristics from the decomposability variables so it is possible to assess the predictors' net effect.

As shown in Table 4.2, different models are estimated in order to compare the significance and magnitude of coefficients. Model 1 tests the direct effect of the level of decomposability of the knowledge base (*KB*) and the collaborative network (*SN*) on the intellectual capacity (*IC*). As it can be seen, while the relation between the decomposability of the *KB* is not significant statistically, the relation with regard to the decomposability of the *SN* is positive and significant in direct support of Hypothesis 7.

Dependent variable: Intellectual capacity (Unit of observation: firm-year)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
constant	0.0338 (0.021)	0.0623 (0.021) **	0.0968 (0.024) ***	0.1101 (0.017) ***	0.1166 (0.035) **	0.0792 (0.029) **	0.0403 (0.027)
patents (log)	0.0682 (0.002) ***	0.0669 (0.002) ***	0.0645 (0.002) ***	0.0604 (0.002) ***	0.0649 (0.003) ***	0.0656 (0.002) ***	0.0692 (0.003) ***
R&D intensity	0.0188 (0.025)	0.0145 (0.025)	0.0131 (0.025)		0.0108 (0.026)	0.0249 (0.02)	0.0376 (0.02) +
number of employes (log)	-0.0007 (0.005)	-0.0007 (0.005)	0.0005 (0.005)		0.0005 (0.005)	-0.0064 (0.003) *	-0.009 (0.002) ***
return on assets	0.0001 (0)	0.0001 (0)	0.0001 (0)		0.0001 (0) +	0.0001 (0) *	0.0002 (0) **
KB's size (log)	-0.0049 (0.007)	-0.0099 (0.007)	-0.0289 (0.009) **	-0.0297 (0.007) ***	-0.0296 (0.012) *	-0.0224 (0.01) *	-0.0144 (0.01)
Inventors by class	-0.0039 (0.006)	-0.0057 (0.006)	-0.0089 (0.007)	-0.0093 (0.006) +	-0.0144 (0.009) +	-0.0079 (0.007)	-0.0097 (0.006)
Years patenting	0.0031 (0.001) ***	0.0027 (0.001) ***	0.0027 (0.001) ***	0.0023 (0) ***	0.0026 (0.001) ***	0.0024 (0.001) ***	0.0026 (0) ***
Control mean level of K	0.141 (0.03) ***	0.1464 (0.03) ***	0.1545 (0.03) ***	0.1548 (0.023) ***	0.1547 (0.03) ***	0.1471 (0.025) ***	0.1573 (0.022) ***
Decomp of KB	-0.0187 (0.026)	-0.0745 (0.029) *	-0.0342 (0.031)	-0.0227 (0.024)	-0.037 (0.034)	-0.0307 (0.031)	-0.03 (0.031)
Decomp of SN	-0.0009 (0.022)	-0.0943 (0.031) **	-0.0337 (0.035)	-0.0439 (0.028)	-0.0576 (0.038)	-0.0309 (0.035)	-0.0008 (0.035)
interaction (Decomp of KB) (Decomp of SN)		0.2677 (0.062) ***	0.1538 (0.067) *	0.1686 (0.053) **	0.1964 (0.075) **	0.1574 (0.067) *	0.0923 (0.065)
interaction (Decomp of KB) size KB			0.0007 (0) ***	0.0007 (0) ***	0.0007 (0) ***	0.0007 (0) ***	0.0007 (0) ***
interaction (Decomp of SN) size SN			-0.0001 (0) *	-0.0001 (0) *	0 (0)	-0.0001 (0)	-0.0001 (0) +
KB's clusterization					-0.0218 (0.022)	-0.0222 (0.02)	-0.0182 (0.02)
KB's density					0.0432 (0.048)	0.0453 (0.044)	0.0489 (0.045)
KB's number of blocks					-0.0006 (0.001)	-0.0014 (0.001) +	-0.0029 (0.001) ***
SN's clusterization					0.0239 (0.017)	0.0233 (0.015)	0.0274 (0.015) +
SN's density					-0.0509 (0.035)	-0.031 (0.03)	-0.0277 (0.029)
SN's number of blocks					-0.0002 (0)	-0.0001 (0)	0.0002 (0)
SN's mean tie					-0.0013 (0.007)	0.003 (0.006)	0.0059 (0.005)
Number observations	1704	1704	1704	2362	1647	1647	1647
Number groups	220	220	220	263	206	206	
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
R-sq overall	0.5906	0.5925	0.5988	0.5591	0.5996	0.6151	0.6195
within	0.4189	0.4262	0.4332	0.4130	0.4430	0.4404	
between	0.6008	0.6113	0.6158	0.5769	0.6179	0.6543	

(Standard Error) *** $p<0.001$ ** $p<0.01$ *$p<0.05$ +$p<0.1$

Table 4.2: Panel regression with firm fixed effects
Descriptive statistics of variables showed in Table B.7 in the Appendix.

Model 2 includes the interaction between the two structures of the firm: the *KB* and the *SN*. This model tests directly Hypothesis 9 but it also provides an interesting insight on the analysis. The estimated coefficient is significant and strongly positive. Considering that the entire range of the decomposability levels is the unity as well as the one of the *IC*, an estimated coefficient of 0.46 indicates a strong impact in the expected value. However, both coefficients for the levels of decomposability have negative signs. This does not contradict Hypothesis 5 and 7 but it demands a more comprehensive interpretation. The negative coefficients of the knowledge and social decomposability when including the interaction must be read as the effect on the *IC* when the other decomposability is null. Therefore, negative signs suggest that the *IC* of the firm depends on both dimensions being simultaneously decomposable. High and similar decomposability levels are associated with a great capacity of the firm for dealing with complex knowledge. Decomposing only the *KB* but having a non-decomposable *SN* (or the symmetric case) is associated with lower *IC*. These results support Hypothesis 5, 7 and 9.

The estimation of Model 3 includes the interaction of the decomposability with the size of both the knowledge base and the collaboration network. The estimated coefficient for the interaction between the two level of decomposability continues to be statistically significant and positive. Regarding the new variables, the interaction between the size of the knowledge base and its level of decomposability has a significant and positive effect on the firm's intellectual capacity for innovating. Considering the entire range of variation of the knowledge base's size (Table B.7), the role a decomposability plays a major role in affecting the capacity for innovating of the firm. Although the empirical analysis supports Hypothesis 6, it does not with Hypothesis 8. The interaction between the size of the collaborative network and the level of social decomposability turns out to be negative. However, the magnitude of the coefficient is small even considering its wider range of variation.

Model 4 and 5 are run in order to test the robustness of previous results. Model 4 rules out 3 variables that reduces the size of the data set since they are not available for all observations. Model 5, on the other hand, includes other variables that might describe the *KB*'s structure and the *SN*'s. On both models, the magnitudes of estimated coefficients for the effect of decomposability remain largely unchanged.

If the logic of Panel with fixed effects is questioned, Model 6 and 7 drop this assumption. Model 6 repeats the statistical analysis of Model 5 but using random effects. As it can be observed in Table 4.2, estimated coefficient do not significantly change either in sign or magnitude. Regarding Model 7, this does not use a Panel. Since there are variables that describe firms and control those characteristic that may influence their

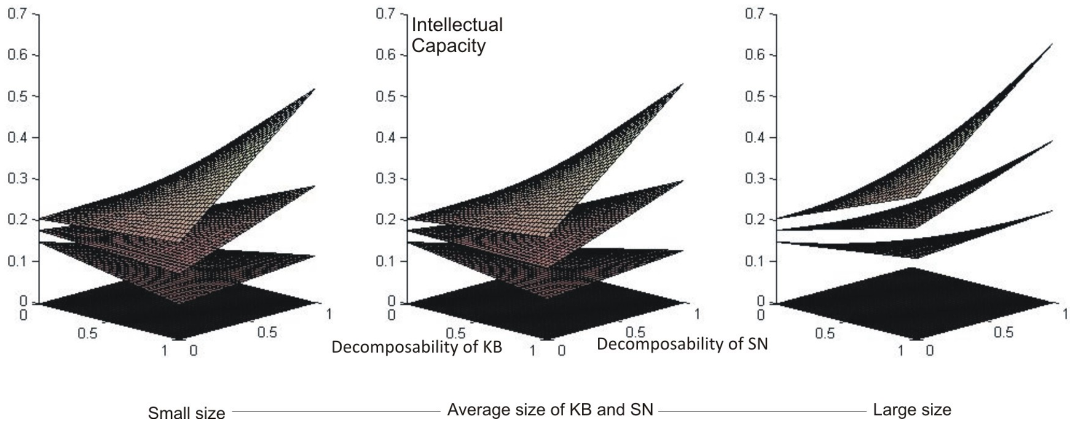


Figure 4.6: Model 3 from Table 4.2 plotted for 3 levels of size of KB and SN. The figure in the middle plots for the median value of KB and SN while figure in the left plots the estimated surface for the 25^{th} percentil and figure in the right for the 75^{th} percentil of their respective distributions. The 3 surfaces in each figure depict the estimated coefficients and their 95% confidence intervals.

intellectual capacity, characteristics that change across firms and across time, Model 7 uses a simple regression analysis. As expected, the number of employees, *R&D* intensity, performance, and size of the knowledge base play a significant role in explaining the *IC* since dummies for firm were dropped. Once again, the sign and magnitude of explanatory variables do not significantly change.

Jointly it can be found empirical evidence to support Hypothesis 5, 6, 7 and 9. The capacity of organization for innovating with complex technologies does seem to be strongly associated with the intrinsic ability of decomposing knowledge into simple structures, an ability that becomes more important the larger the knowledge base is.

4.7 Robustness analysis

The second statistical analysis seeks corroborating results and directly testing Hypothesis 10. The different empirical approach tries analyzing the relation between the success of very complex innovations and firms' capacity for decomposing technology. By dropping the explicit measurement of *IC* the analysis is intended to measure the direct impact of decomposing knowledge when dealing with different levels of complexity. Therefore, the controversy that measuring *intellectual capacity* may generate can be left apart. However, the idea of a global capacity of firms for solving high complex problems that is manifested when innovating is lost.

Differently to the previous analysis, the unit of observation is changed from the firm to the patented invention. The range of inventions' interdependence K is not circumscribed to the unity as it was done for calculating the *IC* (Section 4.4.3) but it can take any value. Furthermore this approach explains the variation in a patent's citation rate in terms of the number of components (N) and the interdependence among them (K) as Fleming and Sorenson (2001) proposed. $N$is included as $1/N$ and $1/N^2$. K, on the other hand, is included both linearly and squared. The ratio among them is also considered.

Following this approach, the model includes controls for *prior art citation*, *number of main technological classes*, *repeated trials* and a *dummy variable for single-class* patents. All them as calculated as Section 4.4.3 previously explained. Furthermore, controls for *number of authors*, *previous combinations used* in the firm and *subclasses already used* by the firm are also considered. All these measurements are explained also at Section 4.4.3.

All the previous variables are specific to each patent. Regarding the firm, a large set of variables is incorporated. The main explanatory variable of a patent's success is the decomposability of its knowledge base as well as the collaboration network's. As it was done in the previous analysis, the level of decomposability is included interacting with the size of the sets according to Hypothesis 6 and 8. The interaction of both measurement among them is included as well.

Differently to the first empirical analysis, as it was explained before, in this one I also include the interaction of the decomposability with the level of interdependence of the patent (K). Since this parameter depicts the roughness of the technological landscape where the patent was created and since I have chosen it to describe the intellectual difficulty of coming up with a successful invention, it would be expected that the larger it is, the less success the patents would have.

Controls for R&D investments relatively to the size of the firm (*R&D intensity*), the numbers of years the firm has been patenting (*age*), the size of the firm in terms of *employees* and its *performance* measured as return on assets are included.

Regarding its knowledge base, its size (*number of sub-classes*) and the ratio between inventors and the the size (*inventors by class*) are considered as controls. Furthermore, other measurements of structural characteristics of the firm are considered. For both the knowledge and inventor's network, the level of *clusterization*, the number of *unconnected blocks*, and the *density* are calculated for controlling different characteristics beyond its level of decomposability. For the network of inventors is also calculated the *mean intensity of ties*.

A special control is also introduced: the *decomposability index proposed by Yayavaram and Ahuja (2008)* as an alternative way of measuring decomposability. Yayavaram and Ahuja (2008)'s propose using a modified version of a standard cluster coefficient for networks. While the standard measurement averages the density of the sub-networks associated to each node, their decomposability index proceeds similarly but considering the density of strong ties in sub-networks. Strong and weak ties are empirically distinguished by comparing the coupling intensity of a firm's knowledge base with the median value of the industry controlled by year and knowledge base's size. The final measurement ranges between 0 and 1, being 1 a fully integrated network and zero a fully decomposable one. Nearly decomposable structure would lie in between those two cases and they would be associated with the best performance on innovating. That is why, this measurement is considered as a controlled with squared effects. Since this index is modified both in the scale and its calculation, the way of calculating it is fully disclosed in Section B.2 in the Appendix. The model incorporates this measurement in order to test whether it explains better the variation of citations.

Following Fleming and Sorenson (2004), a very specific control is included: *scientific references*. In this publication, Fleming and Sorenson proposed that when dealing with high interdependent technologies, the direct consultation of scientific research positively impacts in the success of inventions as it provides "some understanding of the underlying landscape". In order to consider this competing mechanism for explaining success for different levels of K, an algorithm consulted for all patents included in this research directly from the USPTO on-line database counting how many citations they have as "Other references" (non-patent backward citation or reference). The scope of this paper does not allow for the strict classification in scientific index journals, non-index journals, corporate non-technical, corporate technical, book, technical report and conference proceedings references. Since Fleming and Sorenson (2004) reported that over 67% of "Other references" were listed as scientific journals as well as 81% of patents with non-patent references included scientific work, the entire set of non-patent references were considered for each patent. Lastly, following Fleming and Sorenson (2004), the variable is constructed as a dummy variable where 1 indicates the presence of "Other references" in the patent, and 0 otherwise.

Finally, the whole model uses dummy variables for differentiating between *firms*, *publication year* and *main technological class* effects. The control for firm attempts capturing the different propensity to be cited according the creator of the innovation. First, there are several reason for expecting a patent to be more cited by patents of the same firm. Firms enhance their own inventions and therefore they will cite them; patents, sometimes, can be also pieces of bigger inventions (families of patents) and hence they have to cite the rest of pieces; or simply firms are more prone to reference their own patents not only because they know them better but also because it does not represent legal conflicts. Second, since some firms in the data base have developed considerable large numbers of patents compared with other, this fact by itself makes those patents be prone to be much more cited than patents that come from firms with smaller patent bases. Furthermore, this variable matters because of the nature of the data base of this paper where only 300 firms are considered, with a highly asymmetrical distribution of large number of patents (more than 100.000 patents).

4.7.1 Statistical Analysis and Results

Differently to the first statistical analysis, now hypotheses are tested by using a general negative binomial regression model (GNBR). Instead of using the previously developed concept of technical intelligence, forward citation of patents is used a dependent variable as it proxies its technical and commercial success. Due to its nature, a negative binomial model is chosen as it is a count variable and there is not reason *a priori* for assuming that its mean and variance are the same. Furthermore, since this analysis is based on Fleming and Sorenson (2001)'s approach, a generalized version of the model is also considered in order not to assume the variance as constant and independent from the set of independent variables.

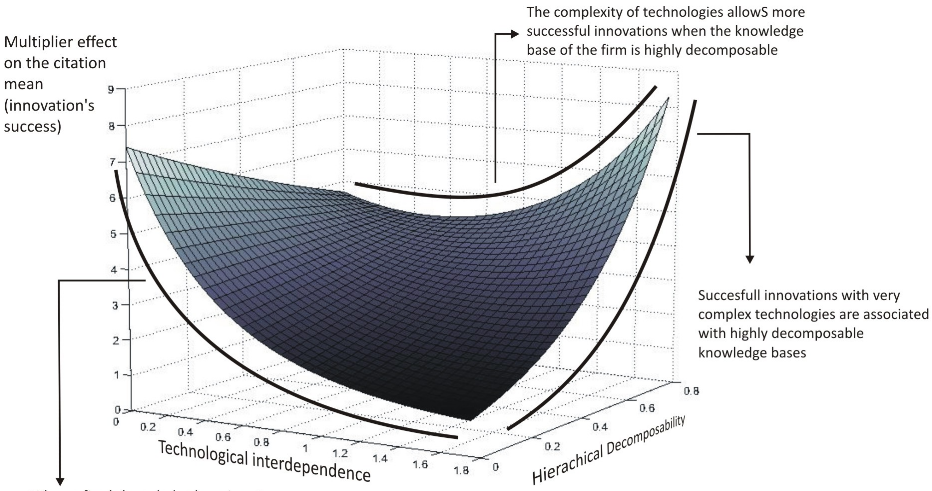
With this new use of the data set, a Panel is not longer suitable for the analysis since for each year and firm there are several observations. Because of this, for controlling different unobserved characteristics of firms and years dummy variables are included. As explained before, all control variables included for developing the measurement of the technical intelligence of the firm, the control variables used for the previous statistical analysis and a new set of controls are included as explanatory variables in this second statistical analysis. The goal is to isolate the variation of the citation rate of patents from other possible sources. The descriptive statistics from this long list of variables are shown in the Section B.4 in the Appendix in Table B.8, while Table B.9 shows linear correlations among them in order to detect possible collinearity. This information and variance inflation factor tests discard the presence of multicollinearity on the data set.

Table 4.7 shows the results of the regression analysis. Since it is a generalized model, coefficients are shown both for the mean and for the variance (for the sake of simplicity all variables included are considered to affect both). 4 models are estimated in order to

General Negative Binomial model citation counts

Variables / Models	MEAN									
	Model 1	**Model 2**	**Model 3**	**Model 4**	**Model 5**	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
constant	2,7792 (0,317) ***	5,9513 (0,961) ***	4,399 (1,159) ***	4,9778 (1,347) ***	5,0903 (1,194) ***	-1,2875 (0,493) **	0,2365 (1,355)	-2,0802 (1,594)	-0,7969 (1,873)	-2,8131 (1,702) +
prior art	0,0053 (0) ***	0,0053 (0) ***	0,0054 (0) ***	0,0055 (0) ***	0,0053 (0) ***	-0,0003 (0)	-0,0004 (0)	-0,0003 (0)	-0,0004 (0)	-0,0005 (0)
dummy single class	-0,4026 (0,148) **	-0,4108 (0,148) **	-0,4103 (0,148) **	-0,3296 (0,151) *	-0,3579 (0,149) *	0,1388 (0,194)	0,1331 (0,194)	0,1407 (0,193)	0,1028 (0,198)	0,0979 (0,196)
Number of main classes	0,0358 (0,008) ***	0,0358 (0,008) ***	0,036 (0,008) ***	0,0385 (0,008) ***	0,0379 (0,008) ***	-0,0014 (0,011)	-0,0018 (0,011)	-0,0016 (0,011)	0 (0,011)	-0,0004 (0,011)
Repeated trials control	-0,0003 (0) +	-0,0003 (0) +	-0,0003 (0) +	-0,0003 (0)	-0,0003 (0)	-0,0008 (0) **	-0,0008 (0) **	-0,0008 (0) **	-0,0008 (0) **	-0,0009 (0) **
dummy sctf ref x K	0,6567 (0,041) ***	0,6562 (0,041) ***	0,6584 (0,041) ***	0,6568 (0,042) ***	0,6521 (0,042) ***	-0,1202 (0,058) *	-0,1212 (0,058) *	-0,1161 (0,058) *	-0,0936 (0,059)	-0,09 (0,059)
1/N	-1,3995 (0,228) ***	-1,4146 (0,228) ***	-1,4106 (0,227) ***	-1,2746 (0,232) ***	-1,3332 (0,229) ***	0,5385 (0,293) +	0,5267 (0,292) +	-1,9501 (0,503) ***	-1,9008 (0,536) ***	0,5201 (0,296) +
$1/N^2$	1,3658 (0,339) ***	1,382 (0,339) ***	1,3803 (0,339) ***	1,1708 (0,346) **	1,2625 (0,341) ***	-0,5307 (0,438)	-0,52 (0,438)	0,1383 (0,037) ***	0,1307 (0,044) **	-0,4409 (0,442)
K	-1,0084 (0,195) ***	-2,2408 (0,624) ***	-2,4056 (0,63) ***	-3,5449 (0,756) ***	-3,0335 (0,731) ***	0,0165 (0,325)	-1,6526 (0,891) +	-0,0009 (0,001)	-0,0009 (0,001)	-1,1629 (1,032)
K^2	0,4606 (0,172) **	0,4325 (0,163) **	0,4427 (0,17) **	0,4739 (0,168) **	0,4878 (0,17) **	-0,4531 (0,4)	-0,5012 (0,369)	0,5416 (0,292) +	0,5161 (0,3) +	-0,6878 (0,513)
K/N	0,2089 (0,227)	0,2267 (0,228)	0,2158 (0,227)	0,3087 (0,236)	0,244 (0,234)	0,7139 (0,33) *	0,7289 (0,329) *	-0,5332 (0,437)	-0,4473 (0,448)	0,5973 (0,344) +
number of authors	0,0545 (0,004) ***	0,0543 (0,004) ***	0,0535 (0,003) ***	0,0534 (0,004) ***	0,0543 (0,004) ***	-0,0085 (0,004) +	-0,0088 (0,004) *	-0,009 (0,004) *	-0,0117 (0,004) **	-0,0097 (0,004) *
previous combinations used	-0,0244 (0,015) +	-0,0235 (0,015)	-0,0239 (0,015)	-0,0221 (0,015)	-0,0222 (0,015)	-0,0299 (0,019)	-0,0294 (0,019)	-0,0311 (0,019)	-0,0323 (0,02)	-0,0296 (0,02)
subclasses used	0,136 (0,019) ***	0,1328 (0,019) ***	0,1352 (0,019) ***	0,1297 (0,019) ***	0,1374 (0,019) ***	-0,188 (0,025) ***	-0,1908 (0,025) ***	-0,186 (0,025) ***	-0,1881 (0,026) ***	-0,1835 (0,025) ***
size KB (log)	-0,283 (0,041) ***	-0,7701 (0,158) ***	-0,6019 (0,195) **	-0,6557 (0,239) **	-0,7512 (0,214) ***	0,0818 (0,055)	-0,0755 (0,213)	0,2107 (0,257)	-0,3748 (0,311)	-0,0903 (0,283)
inventors by class	0,0915 (0,015) ***	0,0757 (0,016) ***	0,093 (0,024) ***	0,1408 (0,045) **	0,096 (0,044) *	-0,02 (0,021)	-0,0254 (0,022)	-0,0748 (0,032) *	-0,0107 (0,061)	0,0579 (0,058)
age	-0,0207 (0,007) **	-0,0186 (0,007) **	-0,0236 (0,007) **	-0,0229 (0,007) **	-0,0181 (0,007) **	0,0454 (0,008) ***	0,0448 (0,009) ***	0,0363 (0,009) ***	0,0399 (0,01) ***	0,0237 (0,009) **
R&D intensity	-0,217 (0,36)	-0,2818 (0,369)	0,2563 (0,384)	0,0692 (0,411)		-1,8768 (0,477) ***	-2,021 (0,484) ***	-1,8613 (0,891) *	-1,2067 (1,058)	
Number of employees (log)	-0,1783 (0,026) ***	-0,1666 (0,026) ***	-0,1532 (0,028) ***	-0,1673 (0,032) ***		0,1325 (0,035) ***	0,1343 (0,035) ***	-0,4656 (0,355)	-0,7053 (0,516)	
Performance	0,0005 (0)	0,0004 (0)	0,0004 (0)	0,0005 (0,001)		-0,0011 (0,001)	-0,0011 (0,001)	0,6867 (0,326) *	0,5971 (0,348) +	
Decomposability		-5,0414 (1,457) **	-5,2069 (1,585) **	-4,3065 (1,838) *	-3,7345 (1,601) *		-2,2975 (1,997)	-1,3176 (2,129)	-2,4651 (2,46)	-1,1354 (2,182)
Decomp x K		1,8614 (0,888) *	2,0997 (0,884) *	1,8053 (1,016) +	1,3086 (0,981)		2,5276 (1,291) +	2,797 (1,285) *	4,4417 (1,308) **	4,4082 (1,27) **
Decomp x size KB		0,7602 (0,238) **	0,7703 (0,259) **	0,9869 (0,324) **	0,7639 (0,29) **		0,2404 (0,321)	0,0633 (0,344)	0,7888 (0,42) +	0,6585 (0,381) +
Social Decomp				2,5378 (1,111) *	1,8579 (1,034) +				5,1934 (1,543) **	6,3149 (1,449) ***
Social Decomp x K				1,832 (0,663) **	1,6138 (0,653) *				-2,1191 (0,938) *	-2,2089 (0,921) *
Social Decomp x nº inventors				-0,1624 (0,143)	-0,1407 (0,138)				-0,2539 (0,195)	-0,3822 (0,184) *
Decomp x Social Decomp				-3,1135 (1,12) **	-1,7781 (1,068) +				-5,1693 (1,597) **	-5,8506 (1,543) ***
Ahuja's ND			0,5095 (0,854)					1,7397 (1,163)		
Ahujas's ND squared			0,9492 (4,407)					2,6736 (6,14)		
KB's Clusterization			0,0039 (0,278)	-0,2117 (0,305)	-0,2387 (0,278)			-0,0535 (0,382)	-0,1767 (0,426)	0,0962 (0,39)
KB's Density			1,2277 (2,786)	0,4521 (3,105)	-1,5009 (2,983)			12,1919 (4,121) **	5,8386 (4,802)	8,6221 (4,604) +
KB's blocks			0,0026 (0,001) +	0,0029 (0,002) +	0,0035 (0,002) *			0,0021 (0,002)	0,003 (0,002)	0,0033 (0,002) +
SN's blocks			-0,0002 (0)	-0,0002 (0)	-0,0001 (0)			0,0003 (0)	0,0004 (0)	0,0004 (0) +
SN's density			15,824 (4,825) **	12,3908 (6,499) +	10,2561 (6,019) +			-7,1946 (6,526)	-25,8789 (8,822) **	-30,6346 (8,242) ***
SN's clusterization			0,9103 (0,262) **	1,1447 (0,317) ***	1,4954 (0,291) ***			0,5793 (0,362)	1,0601 (0,431) *	1,2926 (0,399) **
SN's mean tie			0,0311 (0,021)	0,0445 (0,024) +	0,057 (0,023) *			-0,0267 (0,028)	-0,0031 (0,031)	0,0035 (0,029)
Dummy Controls for main technological class / firm / granting year										
Num observations	73717	73717	73717	70728	72744					
Prob > chi2	0.0000	0.0000	0.0000	0.0000	0.0000					
Log pseudolikelihood	-236852.09	-236836.67	-236782.62	-227284.77	-233302.87					

Figure 4.7: Results for a general negative binomial regression analysis for number of citations



$$\mu_i = exp\Big(-5.04\ d(KB)_i + 1.81\ d(KB)_i\ (K_i) + 0.76\ d(KB)_i\ mean(log(size(KB)))$$
$$-2.24\ K_i + 0.43\ K_i^2 + 0.23\ K_i/mean(N) + 0.66\ K_i * mean(dummy\ scientific\ ref)\Big)$$

Figure 4.8: Expected mean multiplier effect as a function of K and the level of decomposability of the knowledge base - Model 2.

test the robustness of results when including structural variables of both the knowledge and collaboration networks. Model 1 mimics the approach of Fleming and Sorenson (2001) although it has a different data set and it includes controls for firms. The estimated coefficients, however, show being very similar to those estimated by them except for the interdependence parameter (both linear and squared). While Fleming and Sorenson finds a inverted U shaped in the relation between the citation mean and the K parameter, these results support a decreasing relation between both variables. The result is consistent with the assumption of this analysis: the more complex the technology, the less expected success of an invention[5].

Model 4 includes the whole set of variables while Model 2 and 3 focus on partial inclusions. Model 3 includes the decomposability measurement proposed by Yayavaram and Ahuja (2008) to test whether it explains better the variation of dependent variable than the decomposability index here proposed. As it can be observed, both coefficients are not statistically significant. Furthermore, all controls included for depicting structural variables of both the network of technological classes that describes the knowledge base and the network of inventors do not rest significance to the decomposability measurement. Model 3 supports hypothesis 5, 6 and 10. Even though both the estimated

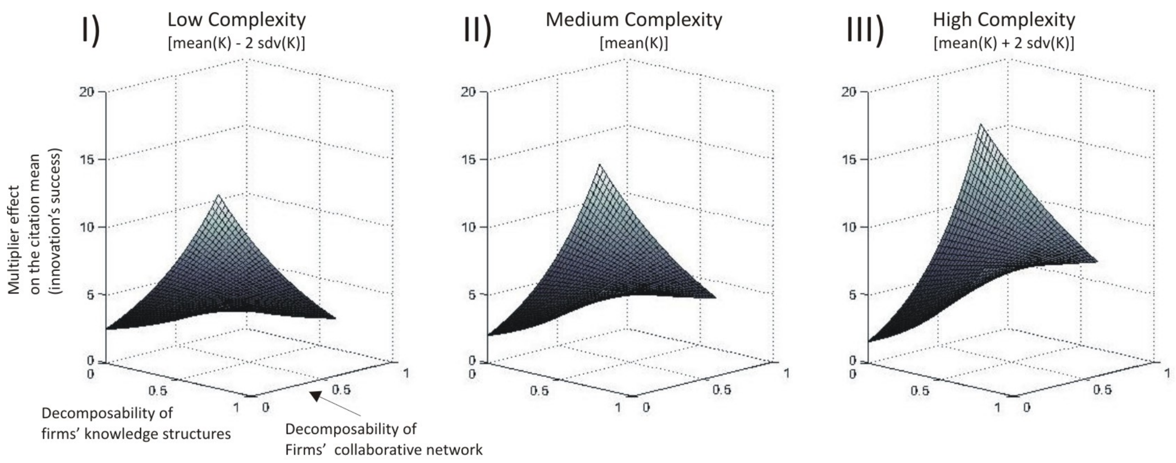
[5]As tested but not shown in this paper, the lack of empirical support for the U-shaped relation is due to the inclusion of firm's characteristics which would indicate that when controlling for the individual more complexity implies less chances of success. The strict replication of Fleming and Sorenson (2001) with this data set *does support* the quadratic effect.

effects of K and the knowledge base's complexity are negative, the interaction between them is strongly positive.

Figure 4.8 helps interpreting the estimation. As it can be observed, the relation between the decomposability and the success of an innovation depends on the complexity of the technology. For low levels of complexity, the relation is negative but as the first increases, the relation becomes strongly positive. Furthermore, the level of complexity negatively affects the expected success of an innovation when the knowledge base is zero decomposable but it turns positive for high levels of decomposability. This clearly indicates that decomposing the complexity allows finding the higher optima hidden in the roughed landscape of possible designs.

Model 4 introduces all variables . It supports all hypothesis globally and it shows statistically significant coefficients directly supporting Hhypothesis 10, 7 and 8. However, due to the amount of variables and the interaction effects, it is easier to read results from a graphical representation. Figure 4.9 helps understanding the results by plotting the effect of the level of decomposability of the firm in its knowledge and social structure on the expected citation rate. Since the level of decomposability interacts with the size of both structures, and the plot has a maximum of 3 dimensions, these 2 variables were taken as constant in their mean value (391.11 sub-classes for the knowledge base and 1089.8 inventors for the collaboration network). The same happens with the interdependence parameter K which interacts with the levels of decomposability. Differently to the previous two, the parameter K is considered in three different values: its mean value (Panel II), its mean value minus 2 standard deviations (Panel I) and plus 2 standard deviations (Panel III). Panel I shows the impact on citation rates of different levels of decomposability on both structures for low levels of complexity. Panel II does the same for medium levels while Panel III for high levels of complexity.

In the three Panels, the levels of decomposability in both structures (horizontal axes) positively impact on the citation rate as long as the other one has low levels. However, when both structures are highly decomposable their joint impact on citation rates change according the complexity of technology. For low complex technologies there is a clear inverted U shape: structures too much decomposable (a near-decomposable structure) would not be necessary but it would also be negatively associated with the outcome of innovating. However, when the complexity of technology is high, near-decomposable structures have a stronger impact on the innovation success. The more decomposable are both structures, the higher the associated innovation's success. Furthermore, the more decomposable both structures simultaneously are, the success is



$$\mu_i = exp\Big(-4.31\ d(KB)_i + 1.81\ d(KB)_i\ mean(K) + 0.99\ d(KB)_i\ mean(log(size(KB))) + 2.54\ d(SN)_i + 1.83\ d(SN)_i\ mean(K) - 0.16\ d(SN)_i\ mean(log(size(SN))) - 3.11\ d(KB)_i d(SN)_i - 3.54\ mean(K) + 0.47\ mean(K)^2 + 0.31\ mean(K/N) + 0.66\ mean(K) * mean(dummy\ scientific\ ref)\Big)$$

Figure 4.9: Expected mean multiplier effect as a function of firms' ability to decompose knowledge and the social structure for different levels of complexity.

even higher. Summing up, the innovation activity of firms in the semi-conductor industry registered in patents seems to support the whole set of Hypotheses.

4.8 Discussion

Based on the theoretical framework of Kogut and Zander (1992), this paper proposes analyzing the intellectual capacity that organizations develop throughout their innovations given that they can be measured in their success and their complexity. The division of labor can be one of the greatest evolutive jumps humanity did as social beings since it planted the seed for the vast complexity of societies nowadays. The volume of knowledge we access and use on daily bases largely exceeds our human capacities. It is the social system as a whole the one which stores it, processes it and improves it. Within this extremely intricate social system, organizations are proposed to play a major role in generating knowledge. Markets and anarchic systems in general have serious limitations on the solutions of some complex problems. As a general rule, if a complex problem can be decompose into simpler parts, the more interdependent are these parts, the more needed is a central mechanism of coordination. Organizations, thus, are capable of processing knowledge that outruns the capacity of single person but are not solvable by descentralized mechanisms.

It is common knowledge that the ability to brake the complexity into simpler structures is a defining characteristic of human brains (Arthur, 1994; Kurzweil, 2012). The pattern recognition on phenomena allow us not only dealing with those problems but it also allows the division of labor when collectively processing information. In other words, it is the capacity of deconstructing the complexity that defines our brain and consequently, our society. However, braking the complexity into manageable parts is not enough for dealing with the most complex problems. Depending on the level of interdependence of the system's component they are differently decomposable and even if decomposable, some centralized mechanism must be capable of considering all subsystems simultaneously for coordinating. Organizations, then, are a social form that gathers people into a meta-identity capable of decomposing and coordinating complex problems.

Although being pretty intuitive, this theoretical understanding of organizations poses a great challenge to be empirically assessed. However, the processing capacity generated by an organized social system can be measured in a particular scenario: innovating. Whenever organizations generate knowledge, they reveal what and how they know (its cognitive structure), how its social network is processed (the physical dimension of the processing machinery), how successful is the solution or the idea they have found for a particular problem, and how complex was the problem they face it. Combining all this information, the intellectual capacity of organizations can be observed. Although it can be criticized the construction of the particular measurement of this capacity here proposed, I consider that the underlying idea of observing the processing power of organizations throughout innovations is a strong approach. However, this "observation" is highly mediated by many theoretical interpretation and assumptions.

First of all this paper bases its analysis on patents representing innovations. This approach, although widely used on the innovation literature, is far from being perfectly accurate. Some innovations are not patented, specially those that fail as projects. If the reason of failing is the high complexity of the technology involved, we cannot know this case as it was never registered. Other patents are part of bigger inventions, not only as part of a family of patents, but also because they can be pieces of larger devices that require many patents to be described. Therefore, patents might be not inventions but pieces of larger inventions.

When measuring the complexity of technologies, this paper follows the approach of Fleming and Sorenson (2001). As these authors mentioned, the NK model (Kauffman and Levin, 1987) can be used as a theory of invention that explains the generation of new technologies as a constant recombination and accumulation of knowledge (Fleming and

Sorenson, 2001; Sorenson et al., 2006). Instead of subscribing to this perspective, I use this model for describing the *intellectual difficulty of combining knowledge*. The NK model, as it describes the difficulty of finding optimal combinations according to the characteristics of the components, it offers a way of capturing the intellectual challenge of innovating independently from the success of the combination."Interdependence makes an invention difficult to perfect, it also enables some very useful combinations" (Fleming and Sorenson, 2001).

This framework demands some questioning. First, other approaches could be used for measuring the complexity or difficulty of knowledge. Some approaches are based on structural characteristics, other on revealed capabilities of developers as the one used by Hausmann et al. (2011). The adaptation of the NK model to the technological landscape offers a valuable theoretical framework because it captures many key characteristics of innovation. Not only it traces the often used parallelism between biological evolution and technological evolution, but it also provides other insights as the modularity problem on design (Baldwin and Clark, 2000). This paper uses the NK model as a way of describing the difficulty of innovating *ex ante*. When a device has high interdependence within its components, it is more difficult to find a successful innovation it is design mainly due to ripple effects. The same difficulty is also associated with a larger room for improvements and the consequent possibility of commercially succeeding. This is the essential logic of what common sense understands as intellectual difficulty and it is perfectly captured by this model. As it captures difficulty, it also provides the means for testing intellectual capacities. Those capable of successfully solving difficult problems prove to be highly intelligent.

Notwithstanding the NK model offers a valuable framework for analyzing innovation, the adaptation to this field demands much thought. Basically it requires the definition of an equivalent of genomes of living beings for technology. Fleming and Sorenson (2001) propose considering patents as genomes of inventions, where technological sub-classes are the gens. This approach, although practical, can be easily defied. Defining what is a technological gen may trigger an endless debate (Arthur, 2009).

Not only the definition of gens as technological sub-classes can be discussed but also the measurement of interdependence on patents. Fleming and Sorenson (2001) use the average number of sub-classes a focal class is combined with as the propensity of interacting with others components. This approach assumes that highly combined sub-classes are easy to combine because they do not interact. This idea can be easily challenged. Some sub-classes might have not been combined before for other reasons rather than interaction. Furthermore, the K parameter is calculated as if the inventor only could have chosen among a only those technological sub-classes of the patent in a binary election. However, inventions could be seen as a binary string of length equal to the total number of sub-classes of the USPTO, where only few of them are activated with 1, and the rest is 0; or they could be thought a string of length equal to the number of sub-classes it has, but instead of being binary, each position in the string have many possible states (the number of possible sub-classes). These are only few examples of how can be altered the adaptation of the NK model to the innovation field. Although there are many implications that can be discussed, the goal here is to point out the large set of assumptions needed for using this framework. Despite this criticism, the K parameter calculated as proposed by Fleming and Sorenson (2001) captures anyways a measurement a difficulty for combining. Since it averages the inverted values of the ease to be combined of technological classes involved in a patent, K, calculated as Equation 4.2 shows, is somehow capturing the originality of this particular combination and therefore the distance from standard ideas (i.e. the intellectual novelty and risk). Other approaches could haven been used for calculating Kalthough for the sake of basing the empirical analysis on familiar ground I chose not to use another one. Analyzing simultaneously several new measurements faces the risk of loosing meaning.

The decomposability measurement proposed in this paper may be difficult to interpret. As a system becomes larger (in term of number of components) and denser (in terms of interaction among its components), the degrees of possibility increases as it can generate complex structures, while its degrees of freedom collapse as it become less adaptable. Hierarchies serves to reduce the density of connections and therefore increase the adaptability of the whole system. By reducing interdependence, they allow the system become larger before it collapses under its complexity. This measurement, then, incorporates the concept of hierarchy for analyzing the knowledge base and collaborative network of firms. By combining the pattern of clusterization and the hierarchical structure, it measures the effect of the hierarchical decomposition: the simplification. Then, when applied to the knowledge base it would capture how much an organization reduces the complexity of knowing; and when applied to the social network, it would capture how adaptable and malleable is the brain power working for it[6]. This measure captures how much simplified is an structure as a percentage of its number of components. An alternative for measuring the same idea would consist in considering the number of connections instead of components and studying how much they are simplified.

[6]In the study of brains' neural networks, this concept is called *wiring economy* (Chen et al., 2006; Rivera-Alba et al., 2011)

As the statistical results show, the intellectual ability to innovate in complexity (IC) is closely related to the size of the firm and the complexity. The social and knowledge structures matter the most as the firm is larger and deals with more complex knowledge. This idea is largely debated among large companies. As they grow into large organizations, sometimes composed by thousands of employees, they may gain resources, stability or power, but the scale might threaten their ability to innovate (Quinn, 1985). As the size and interaction increase, the complexity of its own cognitive machinery increases but also the potentiality of functioning far below its full power.

The idea of the intellectual capacity of organizations is tested on Semi-conductor industry. Comparing firms within the same industry dealing with similar technologies is done in order to isolate possible sources of variation that may affect the analysis when they are not identify. Furthermore, this industry, as explained before, is chosen because of the intense innovation and the sophistication of knowledge involved in their state of the art products. However, analyzing only one sector limits the generality of the empirical results.

Statistical results regarding the decomposability of both structures, specially the knowledge base, with respect to the intellectual capacity or the citation rate of a patent strongly support its relation in comparison with other structural characteristics as the modified clusterization index proposed by (Yayavaram and Ahuja, 2008), standard clusterization index, density or blocks measurements. These results can be read in two different ways. The first and more obvious interpretation is that there is a clear association between the degree of decomposability of a knowledge base and the intellectual capacity of firms. The more complex technologies are and the larger the knowledge bases are, the more important to decompose the knowledge base is in order to succeed innovating. The second reading is that organizations works a intellects. The existence of a hierarchical pattern that is associated with its intelligence suggests that organizations inherits the human way of understanding knowledge. If they were mere sets of people who act independently, there would not be any clear patterns or they may simply not affect the success of their innovations.

The results do not pretend to suggest any direct managerial implication. It could be thought that managers should seek decomposing the firm's knowledge bases or collaborative structure by manipulating couplings either of technology or inventors. Even though it may be influenced, the decomposability of these structures describe a rather exogenous phenomenon associated with cognitive processes when facing complexity than a purposely designed organization. However, they may have an indirect managerial implication.

The two empirical analysis performed in this paper are not redundant. While the second seeks to confirm the first results by a more standard procedure within the innovation literature, the first offers a new perspective of evaluating firms. Stock markets evaluate firms according to expected profits and involved risk; managers also evaluate their decisions by assessing their impact on the value of the firm. However, assessing the impact of decisions might be very difficult and biased towards short run effects. How much would influence on future streams of profits changing the CEO, the acquisition of a firm, the closure of an entire section, or outsourcing some components? What would happen if managers could assess the impact of these decisions into the intellectual capacity of the firm and therefore, the ability to predict earnings in the long term? As psychometric are used to predict job performance, for example, a measurement of the intellectual capacity of firms could also predict long run financial performance or survival.

Results also suggest another idea that is not proposed among the hypotheses. Not only they reveal that both the social and cognitive structure of inventors are strongly associated with the intellectual performance of organizations but they also suggest the bases for an *isomorphism* between both. The maximum effect on the ability to innovate with complex technology happens when the knowledge and the social structure are high and similarly decomposable. This could be indicating that the social and knowledge structure should present some coherence, as if they were two manifestation of the same phenomenon: the organization.

Finally, there is an interesting result observed in both models run in Section 4.6 and 4.7.1 show that the size of the knowledge base magnifies the relation between the level of decomposability and the intellectual capacity of the firm or patent's success. In other words, the larger the knowledge base, the more necessary is it to be decomposable if we want to successfully innovate with that. Thus, it could be said that when a firm deal with a narrow knowledge base, i.e. when it is focused on few technologies, there is not need for decomposing but rather the opposite. Curiously, this relation does not hold of the collaborative network but it is negative in some models. This could be related to the fact that the size of the knowledge base (as it is measured in this paper) cannot be expanded indefinitely because of the finite number of technological classes for patents, while the network of inventors does not have any size limitation.

4.9 Conclusion

This paper subscribes to a long research track about the effects of different complex structures into the innovative activity. Either applied to the research community (Mason and Watts, 2012; Uzzi and Spiro, 2005; Watts and Strogatz, 1998), the organization itself (Rivkin and Siggelkow, 2003; Lazer and Friedman, 2007; Fang et al., 2010; Ethiraj and Levinthal, 2004), an organization's knowledge base (Yayavaram and Ahuja, 2008; Yayavaram and Chen, 2013), or the design of an innovation (Baldwin and Clark, 2000), the structural analysis based on the trade-off sparsity versus density, diversity versus concentration, independence versus interdependence, has played a major role in explaining the expected success of innovations. Based on the general framework proposed by Kauffman and Levin (1987) and Simon (2002), this delicate balance between clusterization and sparsity of complex systems looks like it responds to a "universal architecture" that grantees that near-decomposable systems are the best suited to adapt, survive and evolve towards sophistication.

Within this stream of research, this paper innovates in 4 aspects. First, it proposes that the near-decomposability of organizations is related to its intellectual capacity. For that purpose, it is also proposed an explicit measurement of this concept, a way of capturing its "capacity for process, interpret, encode, manipulate, and access information in a purposeful, goal-directed manner, so it can increase its adaptive potential in the environment in which it operates" (Glynn, 1996). Innovations provides the perfect natural experiment to observe this ability since it can be independently measured the intellectual difficulty they pose along with their success, and they play a major role in the firm's ability to survive. Third, this paper also proposes a new measurement for faithfully capturing Simon's idea of hierarchical near-decomposability. Although very intuitive, the sophistication of Simon's concept cannot be easily translated to a cluster index as some researchers have done. Furthermore, since it plays a major role in explaining the property of complex systems specially in the innovation research, it deserves much attention. The index here proposed incorporates information about the amount of layers, how they relate with each other, the number of clusters and the intensity of *within* and *among* relations to calculate its degree of decomposability. Fourth, this paper analyze organizations in their intrinsic double socio-epistemic dimension. Not only corroborates that near-decomposable structures increases the ability to adapt and survive, it also provides evidence to support that near-decomposability affects or defines the organizational intellectual capacity. This paper also suggests that it is the near-decomposability of *both* the social and the knowledge structure of the firm that explain this characteristic. Furthermore, the decomposability of both structures maximize the intellectual capacity of organizations not only when they have high levels of decomposability but also when both the knowledge and the social structure do not differ in those levels. This may indicate the very intimate nature of these two sides of the sophisticated cognitive machinery an organization is. The hierarchical near-decomposability of the internal structure of socio-epistemic organisms determine their level of intelligence and ultimately its ability to survive and success in its environment. Finally, oppositely to Fleming and Sorenson (2001), I find that the relation between complexity and success is not necessarily negative for high levels but it can be positive if the innovator is sophisticated enough. This assertion is pretty intuitive. Creating a spaceship capable of landing in the moon back in 1969 was for sure an enterprise as complex as successful. Complexity is associated with success as long as you are capable of dealing with it.

Appendix A

Appendix to Chapter 3

A.1 Measuring redundancies

This section offers a detailed explanation and illustration of the methodology used in this chapter for describing the structure of organization.

A.1.1 Social distances

An inventor is considered as the set of patents authored by her/him, on a period of three years for the same firm. Inventor I, working for firm f during T is defined as:

$$I_{f,T} = \{p_{f,t} / \; T - 3 < t \leq T \wedge I \text{ is author of } p_{f,t}\} \tag{A.1}$$

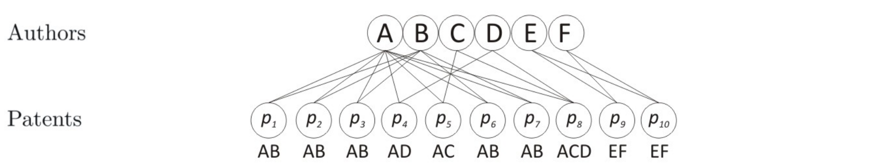


Figure A.1: Example 1

Using the example provided by Figure A.1, inventor has authored patents p_9 and p_{10} ($E = \{p_9, p_{10}\}$), while inventor has authored p_5 and p_8 ($C = \{p_5, p_8\}$) . In order the depict the co-authorship network, first, the adjacency matrix is defined as the matrix whose element on the i^{th} row and j^{th} column represents the cardinal of the intersection between inventors I_i and I_j, i.e. the number of patents they have co-authored.

$$A_{f,T} = [a_{ij}]_{nxn} \ / \ \ a_{ij} = \#(I_{i,f,T} \cap I_{j,f,T}) \tag{A.2}$$

Co-authorship Matrix $A_{6x6} = \begin{pmatrix} 0 & 5 & 2 & 2 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{pmatrix}$

Figure A.2: Example 2

The adjacency matrix at Figure A.2 follows the same simple example. As it can be observed, for instance, inventor A has co-authored 5 patents with inventor B. This indicates a strong collaboration tie which leads to a short social distance. Thus, once defined the adjacency matrix of the co-authorship network, the matrix of *direct* distances among inventors is stated as a the maximum intensity of the adjacency matrix minus the intensity of each couple (as long this is positive). This way, the more intense a collaboration tie is, the *closer* two inventors are in this space.

$$DD_{f,T} = [dd_{ij}]_{nxn} \ / \ dd_{ij} = \begin{cases} max(A_{f,T}) + 1 - a_{ij} & \text{if } a_{ij} > 0 \\ 0 & \text{if } a_{ij} = 0 \end{cases} \tag{A.3}$$

Direct Distances $DD_{6x6} = \begin{pmatrix} 0 & 1 & 4 & 4 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 5 & 0 & 0 \\ 4 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 4 & 0 \end{pmatrix}$

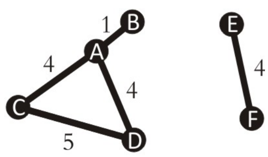


Figure A.3: Example 3

As it can be observed at Figure A.3, inventors A and B are the closest ones since they have the strongest collaboration tie, whose distance is set to 1. Inventors C and D, on other side, have the maximum direct distance since they have the weakest collaboration tie: only 1 patent co-authored. Using direct distances among inventors, the full matrix of distances is calculated afterwards. The distance between any couple of inventors is calculated as the geodesic distance by using the ties and their length as stated in the matrix of direct distances. The geodesic is the length of the shortest path between every couple of inventors. At example, inventors B and D, although not directly connected, have a social distance of length 5 since the shortest path between them is $B - A - D$.

$$D_{f,T} = [d_{ij}]_{nxn}/d_{ij} \text{ is the length of the SP between } i \text{ and } j \tag{A.4}$$

Geodesic Distances $D_{6x6} = \begin{pmatrix} 0 & 1 & 4 & 4 & 0 & 0 \\ 1 & 0 & 5 & 5 & 0 & 0 \\ 4 & 5 & 0 & 5 & 0 & 0 \\ 4 & 5 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 4 & 0 \end{pmatrix}$

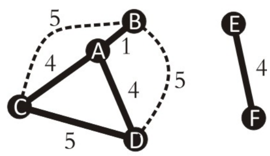


Figure A.4: Example 4

For those cases where there is not any possible path between two inventors (like between inventors E and A), since the distance cannot be defined, it is assumed to be the maximum distance registered in this matrix. The distance between E and A is equal to 5 in the example at Figure A.4. This assumption is based on the fact that all inventors work for the same firm and therefore, they are socially connected. The impossibility of defining the distance between E and A does not mean that they are infinite distant, but it means that their distance is greater that the one observed. Therefore, those undefined distances are assumed to be equal to the maximum registered distance in the network. Finally, the matrix of collaboration distances among inventors of firm f in time T is defined as:

$$D_n = [\delta_{ij}]_n \text{ where } \begin{cases} \delta_{ij} = \text{length of SP between } i \text{ and } j \text{ if there is at least one path between them} \\ \delta_{ij} = max(D_n) \text{ if there is not any path between } i \text{ and } j \end{cases} \tag{A.5}$$

Full Matrix of distances $D_{6x6} = \begin{pmatrix} 0 & 1 & 4 & 4 & 5 & 5 \\ 1 & 0 & 5 & 5 & 5 & 5 \\ 4 & 5 & 0 & 5 & 5 & 5 \\ 4 & 5 & 5 & 0 & 5 & 5 \\ 5 & 5 & 5 & 5 & 0 & 4 \\ 5 & 5 & 5 & 5 & 4 & 0 \end{pmatrix}$

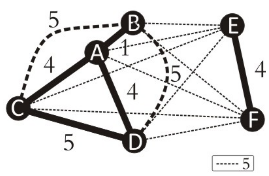


Figure A.5: Example 5

The matrix of social distances is symmetric and positive definite, and it is only composed by integer numbers. The symmetry indicates that the distance from inventor I_i

to inventor I_j is the same that the distance from inventor I_j to inventor I_i. Considering only once every possible couple of inventors, a vector of collaboration distances is built as:

$$SD_{f,T} = [\delta_h]_{n(n-1)/2x1} / \ \delta_h = d_{ij} \ \forall \ (0 < i < n) \ (i < j \leqslant n)(i \neq j) \tag{A.6}$$

Following the example, the vector of social distances would be:

$$SD = [1\ 4\ 4\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 5\ 4]_{1x15}$$

A.1.2 Knowledge distances

For building the knowledge space, I follow a similar process. Now, instead of using inventors, I use technological classes. First, I define a technological class as:

$$C_{i,f,T} = \{p_{f,t} : (T - 3 < t \leq T) \wedge (p_{f,t} \text{ is classified on } C_i)\} \tag{A.7}$$

In this case, patents *couple* technological classes. Then, the adjacency matrix of the coupling network is defined as:

$$B_{f,T} = [b_{ij}]_{mxm} \ : \ b_{ij} = \#(C_{i,f,T} \cap C_{j,f,T}) \tag{A.8}$$

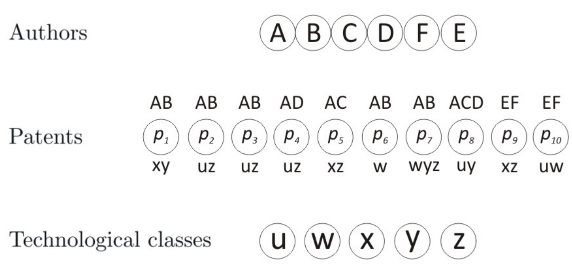
Once defined the adjacency matrix, the matrix of distances is calculated following the same procedure used for the matrix of social distances. First, direct distances are proposed to be the as the maximum intensity of the matrix minus the intensity of link plus 1; then, geodesics are calculated; and finally those distances that are not defined are assumed to be as long as the maximum distances registered in the matrix.

$$D_n^{(k)} = [\delta_{ij}]_{mxm} \text{ where } \begin{cases} \delta_{ij} = \text{length of SP between class } i \text{ and } j \text{ if there is at least 1 path between them} \\ \delta_{ij} = max(D_n) \text{ if there is not any path between class } i \text{ and class } j \end{cases} \tag{A.9}$$

This matrix depicts how different technologies are coupled by certain firm during 3 years. Those technological classes that are *similar* or *close*, are technologies frequently coupled by the firm, meaning they are understood, known or believed to be naturally combinable.

Once this distance matrix is defined, a multidimensional scale is performed in order to generate a set of coordinates for each technological class such that the Euclidean distances between are a monotonic transformation of the corresponding dissimilarities in A.9. The set of coordinates for each technological class is defined as:

$$\Pi_{f,T} = [\pi_{ij}]_{mxr} : \pi_{ij} \text{ is the } j^{th}\text{coordinate of class } i \tag{A.10}$$



$$B_{5x5} = \begin{pmatrix} 0 & 1 & 0 & 1 & 3 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 1 & 1 & 1 & 0 & 0 \\ 3 & 1 & 2 & 0 & 0 \end{pmatrix} DD^B_{5x5} = \begin{pmatrix} 0 & 3 & 0 & 3 & 1 \\ 3 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 3 & 2 \\ 3 & 3 & 3 & 0 & 0 \\ 1 & 3 & 2 & 0 & 0 \end{pmatrix} D^B_{5x5} = \begin{pmatrix} 0 & 3 & 3 & 3 & 1 \\ 3 & 0 & 5 & 3 & 3 \\ 3 & 5 & 0 & 3 & 2 \\ 3 & 3 & 3 & 0 & 4 \\ 1 & 3 & 2 & 4 & 0 \end{pmatrix}$$

$$\Pi_{5x3} = \begin{pmatrix} 0.054 & 0.804 & -1.029 \\ -2.542 & 0.393 & 0.453 \\ 2.364 & -0.588 & 0.373 \\ -0.569 & -2.193 & -0.145 \\ 0.693 & 1.584 & 0.348 \end{pmatrix}$$

Figure A.6: Example 6

In order to provide a simple example, I use the same of Section A.1.1. Now, there are 10 patents at the knowledge base, authored by 6 inventors and classified into 5 technological classes, u, w, x, y and z. Figure A.6 shows the coupling network (B), the direct distances matrix (DD) and the distances matrix (D) for this example. By performing a multidimensional scale on D, Π is obtained, where the 5 technological classes are located in a $\mathbb{R}^3$ space.

The new set of coordinates is then used to depict the knowledge space inventors are located in. As the patent, inventors reveal what they know according what classes their patents are classified in and the frequency of patenting in those classes. Thus, defining the adjacency matrix of the bipartite dichotomous network between the n inventors and p patents that conform the knowledge base of firm f during T as:

$$InvPat_{f,T} = [a_{i,j}]_{nxp} \text{ such that } a_{i,j} = \begin{cases} 1 \text{ if Inventor } i \text{ has authored patent } j \\ 0 \text{ if Inventor } i \text{ has not authored patent } j \end{cases} \tag{A.11}$$

and defining the adjacency matrix of the bipartite dichotomous network between the m technological classes and p patents that conform the knowledge base of firm f during T as:

$$ClsPat_{f,T} = [a_{i,j}]_{mxp} \text{ such that } a_{i,j} = \begin{cases} 1 \text{ if patent } j \text{ is assigned to class } i \\ 0 \text{ if patent } j \text{ is assigned to class } i \end{cases} \tag{A.12}$$

then, it can be observed how inventors patent in different technological classes at matrix $InvClass$:

$$InvCls_{nxm} = InvPat\ ClassPat' \tag{A.13}$$

Following the example, these three matrices would be:

$$InvPat_{6x10} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$$ClsPat_{5x10} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

$$InvCls_{6x5} = \begin{pmatrix} 4 & 2 & 2 & 3 & 5 \\ 2 & 2 & 1 & 2 & 3 \\ 1 & 0 & 1 & 1 & 1 \\ 2 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Figure A.7: Example 7

Matrix $InvCls$ reveals how inventors patent on different classes. Dividing each row of this matrix by its total sum, it is obtained the relative frequency of patenting on different classes of different inventors. I call this matrix W such that:

$$W_{f,T} = [w_{ij}]_{nxm} \text{ such that } w_{ij} = InvCls_{ij}/max(InvCls_i) \tag{A.14}$$

$$W_{6x5} = \begin{pmatrix} 0.250 & 0.125 & 0.125 & 0.188 & 0.313 \\ 0.200 & 0.200 & 0.100 & 0.200 & 0.300 \\ 0.250 & 0.000 & 0.250 & 0.250 & 0.250 \\ 0.500 & 0.000 & 0.000 & 0.250 & 0.250 \\ 0.250 & 0.250 & 0.250 & 0.000 & 0.250 \\ 0.250 & 0.250 & 0.250 & 0.000 & 0.250 \end{pmatrix}$$

Figure A.8: Example 8

W is used as the weights to average the coordinates of each technological class for each inventor. Then, inventors are located in the same space where technological classes according to the frequency of use of those classes. Once located in this space, distances among them are calculated simply as the euclidean one. Finally, a vector of knowledge distances is obtained:

$$KD_{f,T} = [\delta_h]_{n(n-1)/2x1} / \ \delta_h = d_{ij} \ \forall \ (0 < i < n) \ (i < j \leqslant n)(i \neq j) \tag{A.15}$$

In the example, the vector of knowledge distances would be:

$$KD = [0.28\ 0.64\ 0.39\ 0.31\ 0.31\ 0.87\ 0.51\ 0.45\ 0.45\ 0.76\ 0.83\ 0.83\ 0.59\ 0.59\ 0.00]_{1x15}$$

A.1.3 Redundancies

For each firm, each 3-year period (moving quarterly), distances among the n inventors that patent for the firm, both in the social and the knowledge space, are depicted by vectors SD and KD. These two describe the $n(n-1)/2$ relations among inventors in the socio-knowledge space .

Two different transformation are applied on these vectors:

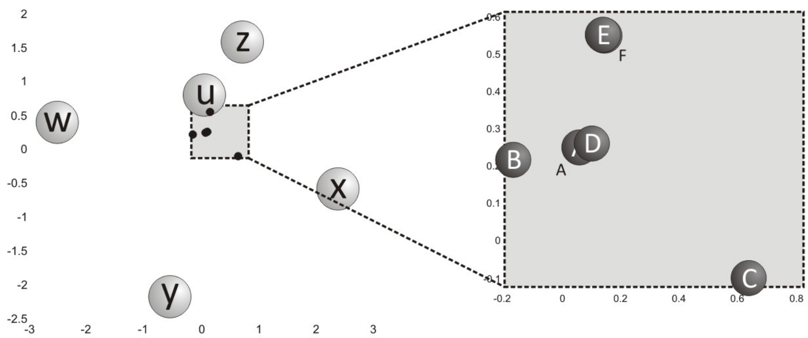


Figure A.9: Example 9

- **Winsorization** In order to discard outliers, values larger than the 99.5 percentile or smaller than 0.5 percentile are forced to be those limits. Then, the whole range of variation of each vector is re-scaled to 1. I call SDw and KDw to SD and KD transformed this way.

- **Standardization** Both vectors, SD and KD are standardized according to their means and standard deviation. I call the transformed vectors SDs and KDs.

Figure A.10 provides a complete example. Panel I and II display a group of 65 inventors working for certain firm. In Panel I they are located according to their social distances, while in Panel II they are dispersed according to their knowledge distances. Figure A.10 shows a clear case of redundancies since there are 3 well delimited social groups (referenced by "*", "+" and "o") and two of them overlap in their research area (groups '+" and "o").

The set of 2,080 distances among inventors at each space is calculated and then plotted at Panel III. In order to facilitate the analysis, distances within groups are plotted using the symbol of the group, while distances inter-group are symbolized with ".", "x" and stars. As it can be observed, distances within groups are shorter than among groups and there is a positive correlation among all of them.

The entire range of variation of distances is then transformed into different ways. First, it is winsorized at 95% of variation, re-scaled to the 1 and plotted at Panel IV. Then, all observations are projected into the uni-dimensional space perpendicular to the space *social distance = knowledge distance* at Panel VI as Equation A.16 shows.

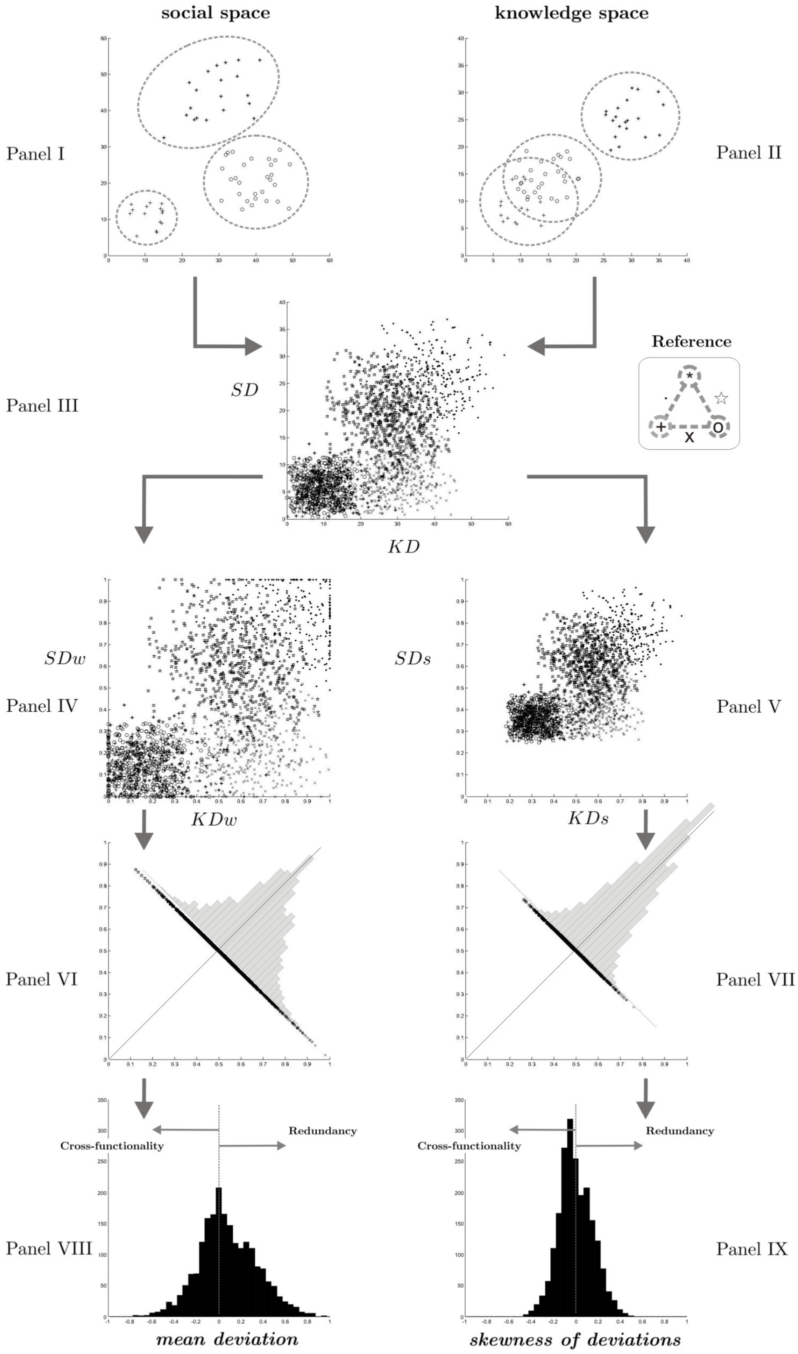


Figure A.10: Illustration of how the two metrics are calculated

$$\begin{aligned} SDw_p(i) &= \frac{(1 + SDw(i) - KDw(i))}{2} \\ KDw_p(i) &= \frac{(1 + KDw(i) - SDw(i))}{2} \end{aligned} \quad \forall\ 1 \leq i \leq \frac{n(n-1)}{2} \tag{A.16}$$

The resulting distribution of observations is re-scaled to the unity and analyzed in its distribution by considered by its mean. Equation A.17 describes the new variable:

$$x_w(i) = \frac{2\left(SDw_p(i)^2 + (1 - KDw_p(i))^2\right)^{\frac{1}{2}}}{\sqrt{2}-1} \forall\, 1 \leq i \leq \frac{n(n-1)}{2} \tag{A.17}$$

Secondly, distances are standardize by considering its mean and standard deviation. Since the distribution is forced to have mean equal to 0 and standard deviation equal to 1, a skewness parameter is chosen for describing the distribution. After standardizing, those observations located further than 3 standard deviations from 1 are forced to be those limits, and then they are projected into the uni-dimensional space perpendicular to the space *social distance* = *knowledge distance* as Equation A.18 describes.

$$\left.\begin{aligned} SDs_p(i) &= \frac{(1 + SDs(i) - KDs(i))}{2} \\ KDs_p(i) &= \frac{(1 + KDs(i) - SDs(i))}{2} \end{aligned}\right. \forall\, 1 \leq i \leq \frac{n(n-1)}{2} \tag{A.18}$$

The new variable after the standardization is called x_s and is obtained as Equation A.19 describes it. Finally, the skewness of the distribution is calculated as the difference between the median and the mean divided by the standard deviation.

$$x_s(i) = \frac{2\left(SDs_p(i)^2 + (1 - KDs_p(i))^2\right)^{\frac{1}{2}}}{\sqrt{2}-1} \forall\, 1 \leq i \leq \frac{n(n-1)}{2} \tag{A.19}$$

Finally, the two measures used in the statistical analysis are:

$$mean\ deviation = mean(x_w) \tag{A.20}$$

$$skewness\ of\ deviations = \frac{mean(x_s) - median(x_s)}{std(x_s)} \tag{A.21}$$

An extra measure was tested but finally discarded due to the high correlation with *mean deviation*. Using the first transformation, and considering the distribution of projected distances in the orthogonal space to the isomorphic line as depicted by Equation

A.17, this metric re-shapes the distribution of x_s, $f(x_s)$, by multiplying it by squared values of x_s keeping the original sign x_s. Since $-1 < x_s < 1$, the larger is x_s, the less affected is the value of $f(x_s)$. The closer to 0 x_s is, i.e. the isomorphic zone, the more $f(x_s)$ is collapsed. Instead of squaring x_p, it is multiplied by its absolute value ($x_p|x_p|$) so it can keep the sign since its indicates whether x_p deviates towards redundancy or cross-functionality. Finally, this measure, is the are comprehend under this new function. It weights more the longer deviations from the isomorphism, and it compares both kind of deviations, negative and positive, for revealing which is the predominant.

$$deviation = \int_{-1}^{1} f(x_p)x_p|x_p| \tag{A.22}$$

Figure A.11 provides a graphic example of how it works. The Panel on the left show the distribution, the central Panel shows in light gray the distribution while in dark gray the transformation, and finally the Panel on the right amplifies the middle Panel for showing how this measure punishes small deviation and weights more longer ones. This measure has a correlation of 98.99% with *mean deviation*. Although its construction differs a lot, it validates what both capture.

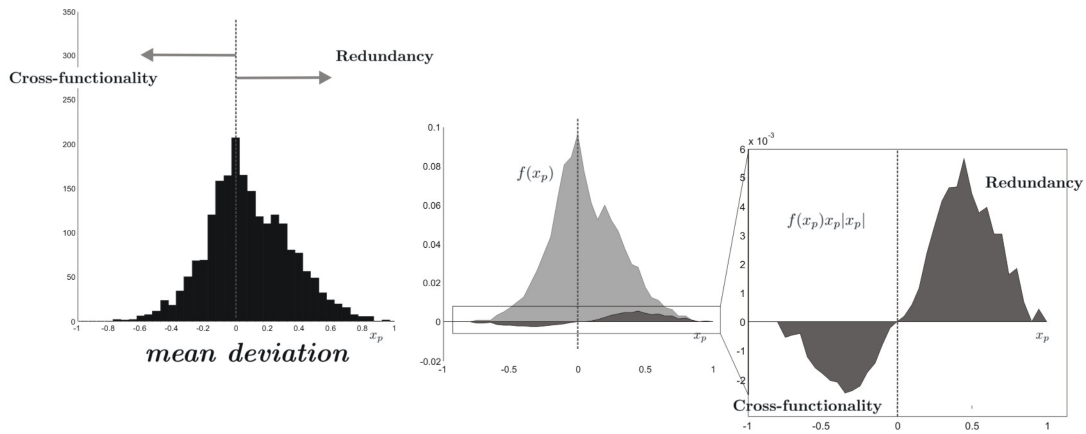


Figure A.11: Weighted density function

A.2 Descriptive statistics

Table A.1: Descriptive Statistics

Level	Variable	Obs	Mean	Std. Dev.	Min	Max
Patent	citation	89016	10.082	15.618	0	313
	prior citations	89016	11.616	18.413	0	568
	dummy single subclass	89016	0.105	0.307	0	1
	number of major classes	89016	1.448	0.716	1	10
	previous trials	89016	6.453	32.051	0	889
	number of authors	89016	2.269	1.460	1	26
	couplings already used by the firm	89016	0.399	0.420	0	1
	sub-classes already used by the firm	89016	0.720	0.351	0	1
	Number of components (N)	89016	3.654	2.359	1	35
	Interdependence (K)	89016	0.229	0.094	.07	4
Knowledge Base	mean deviation	88818	0.265	0.177	-0.446	0.711
	standard deviation of deviations	88818	0.371	0.052	0	1.127
	skewness of deviations	88817	-0.073	0.096	-0.577	0.707
	Inventors / agreggated classes	88704	2.435	1.140	0.610	7.167
	number of years patenting	89016	13.019	5.206	0	24
	Number of inventors (log)	88704	6.502	0.892	4.605	8.203
	KB's Clusterization	89016	0.563	0.045	0	1
	KB's Density	89016	0.026	0.035	0.012	1
	KB's blocks	89016	19.860	7.187	1	46
	SN's blocks	88704	181.937	157.314	6	686
	SN's density	88704	0.009	0.008	0.001	0.062
	SN's clusterization	88704	0.524	0.075	0.267	0.797
	SN's mean tie	88704	1.819	0.574	1.039	3.715
Firm	R&D intensity	87329	0.096	0.054	0.002	1.343
	size of the firm	85382	2.934	1.009	-4.962	4.604
	perfomance	87419	3.414	17.609	-273.034	158.680

Correlation Matrix * p<0.1

		1	2	3	4	5	6	7	8
1	citation	1.0000							
2	prior citations	0.1149*	1.0000						
3	number of major classes	0.0545*	0.0135*	1.0000					
4	previous trials	-0.0351*	-0.0105*	-0.1181*	1.0000				
5	number of authors	0.0605*	0.0539*	-0.0044	0.0016	1.0000			
6	couplings already used by the firm	-0.0177*	0.0633*	-0.2027*	0.2029*	0.0267*	1.0000		
7	sub-classes already used by the firm	0.0129*	0.0641*	-0.1013*	0.0900*	0.0476*	0.5418*	1.0000	
8	Inventors / agreggated classes	-0.0660*	-0.0529*	-0.0083*	0.0508*	0.0718*	0.1134*	0.2912*	1.0000
9	number of years patenting	-0.0356*	0.0307*	-0.0240*	0.0518*	0.0431*	0.1391*	0.2607*	0.5445*
10	R&D intensity	0.0109*	-0.0067*	0.0275*	-0.0066*	-0.0101*	-0.0813*	-0.1367*	-0.0817*
11	size of the firm	-0.0559*	-0.0645*	0.0108*	-0.0000	-0.0003	0.0731*	0.2366*	0.6500*
12	perfomance	0.0328*	-0.0506*	-0.0046	-0.0052	-0.0110*	0.0109*	0.0421*	0.1252*
13	Number of components (N)	0.0603*	0.0421*	0.4463*	-0.1857*	0.0149*	-0.2008*	-0.0168*	-0.0213*
14	Interdependence (K)	-0.0161*	-0.0505*	-0.0914*	0.1375*	-0.0846*	-0.0057*	-0.1438*	-0.0096*
15	mean deviation	-0.0259*	-0.0887*	0.0109*	-0.0235*	-0.0677*	-0.1491*	-0.2260*	-0.0671*
16	standard deviation of deviations	0.0254*	0.0466*	-0.0052*	0.0265*	0.0211*	0.1386*	0.2471*	0.3577*
17	skewness of deviations	0.0167*	0.0790*	-0.0192*	0.0191*	0.0407*	0.1294*	0.1858*	-0.0042
18	Number of inventors (log)	-0.0635*	-0.0237*	-0.0163*	0.0315*	0.0231*	0.1744*	0.4118*	0.7652*
19	KB's Clusterization	0.0416*	0.0616*	0.0793*	-0.0218*	0.0036	0.0447*	0.0708*	-0.1238*
20	KB's Density	0.0679*	0.0527*	0.0505*	-0.0135*	0.0046	-0.0984*	-0.2470*	-0.3478*
21	KB's blocks	-0.0859*	-0.0875*	-0.0166*	0.0181*	-0.0034	0.0591*	0.1964*	0.4751*
22	SN's blocks	-0.0427*	-0.0301*	0.0204*	0.0223*	-0.0072*	0.0769*	0.2324*	0.8552*
23	SN's density	0.0877*	0.0706*	0.0286*	-0.0143*	0.0279*	-0.0884*	-0.2220*	-0.3316*
24	SN's clusterization	0.0215*	0.0038	0.0025	0.0310*	0.1722*	0.0163*	0.0317*	0.2108*
25	SN's mean tie	0.0327*	0.2024*	-0.0320*	0.0109*	-0.0040	0.2073*	0.3028*	-0.0541*

		9	10	11	12	13	14	15	16
9	number of years patenting	1.0000							
10	R&D intensity	-0.0202*	1.0000						
11	size of the firm	0.2354*	-0.2586*	1.0000					
12	perfomance	-0.0049	-0.5384*	0.2257*	1.0000				
13	Number of components (N)	-0.0118*	-0.0337*	0.0092*	-0.0146*	1.0000			
14	Interdependence (K)	-0.0375*	0.1213*	-0.0566*	0.0309*	-0.3109*	1.0000		
15	mean deviation	-0.0191*	0.2070*	-0.0223*	-0.0012	-0.0738*	0.2422*	1.0000	
16	standard deviation of deviations	0.2859*	-0.1266*	0.2067*	0.0267*	0.0274*	-0.1333*	-0.5414*	1.0000
17	skewness of deviations	0.0248*	-0.2043*	-0.0429*	0.0112*	0.0576*	-0.2188*	-0.6694*	0.3545*
18	Number of inventors (log)	0.5371*	-0.1170*	0.8082*	0.1095*	0.0133*	-0.1173*	-0.0935*	0.3176*
19	KB's Clusterization	-0.1188*	-0.0603*	0.0000	-0.0153*	0.0891*	-0.0916*	-0.1401*	0.0072*
20	KB's Density	-0.3378*	0.1366*	-0.5503*	-0.1155*	0.0226*	0.0657*	-0.1117*	-0.0772*
21	KB's blocks	0.2856*	-0.0628*	0.6222*	0.0671*	-0.0007	-0.0896*	0.0094*	0.1679*
22	SN's blocks	0.4356*	0.0009	0.6904*	0.1123*	-0.0088*	-0.0153*	0.0212*	0.3367*
23	SN's density	-0.3052*	0.1878*	-0.5381*	-0.1306*	0.0069*	0.0480*	-0.1636*	-0.0148*
24	SN's clusterization	0.1562*	0.0580*	0.0345*	-0.0820*	0.0076*	-0.0404*	-0.0610*	0.0585*
25	SN's mean tie	0.2144*	-0.0700*	-0.0272*	-0.0732*	0.0514*	-0.1801*	-0.3945*	0.2957*

		17	18	19	20	21	22	23	24
18	Number of inventors (log)	0.0353*	1.0000						
19	KB's Clusterization	0.1238*	-0.0050	1.0000					
20	KB's Density	0.0882*	-0.7170*	0.2169*	1.0000				
21	KB's blocks	-0.0728*	0.6856*	-0.2033*	-0.5360*	1.0000			
22	SN's blocks	-0.0896*	0.7584*	-0.1119*	-0.3630*	0.5633*	1.0000		
23	SN's density	0.1162*	-0.6328*	0.0282*	0.7774*	-0.4298*	-0.2839*	1.0000	
24	SN's clusterization	0.0649*	0.0856*	-0.0052*	0.0263*	0.0069*	0.0119*	0.1745*	1.0000
25	SN's mean tie	0.3445*	0.2492*	0.1685*	-0.1561*	-0.0132*	-0.0162*	-0.0316*	-0.0427*

Figure A.12

A.3 Statistical results with another measure

As it can be observed, these plots present the same behavior of the estimated function as the one presented in Figure 3.6. Since the variable *skewness of deviations* captures

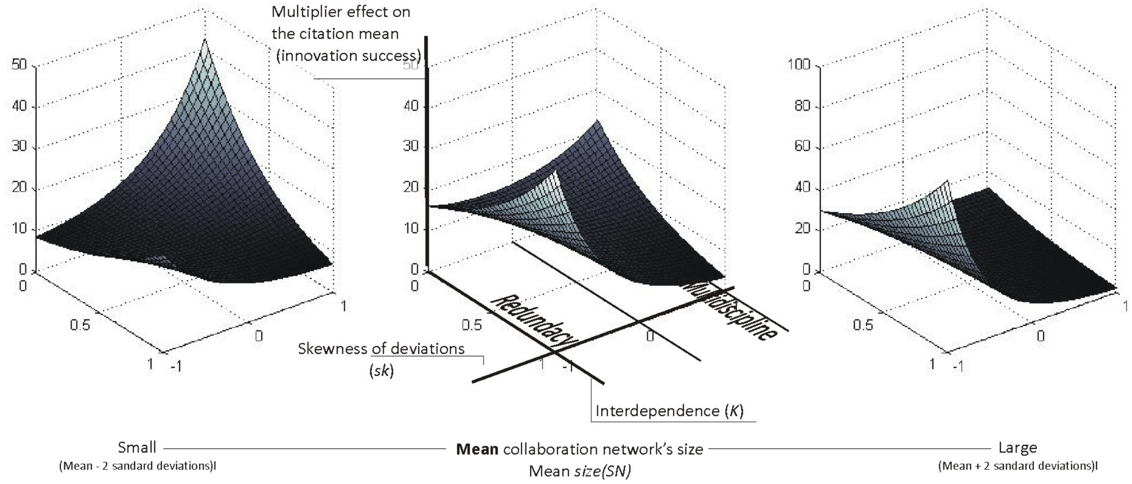


Figure A.13: Expected mean multiplier effect as a function of skewness of deviations (*sk*) and interdependence (*K*) estimated from Model 1.

with negative values the presence of organizational redundancies and with positive values cross-functional structures, then, the plot is inverted regarding the previously mentioned.

Figure A.14: Generalized negative binomial estimates of citation counts (5-year window, standard errors in parentheses)

General Negative Binomial model citation counts

Variables / Models	MEAN Model 1	MEAN Model 2	MEAN Model 3	VAR Model 1	VAR Model 2	VAR Model 3
constant	2.6684 (0.206) ***	2.7563 (0.254) ***	1.2703 (0.204) ***	0.1248 (0.313)	0.0631 (0.379)	0.2599 (0.289)
prior citations	0.0057 (0) ***	0.0057 (0) ***	0.0078 (0) ***	-0.0002 (0)	-0.0002 (0)	0.0001 (0)
dummy single subclass	-0.2897 (0.13) *	-0.2905 (0.13) *	-0.2274 (0.134) +	0.2875 (0.176)	0.2873 (0.176)	0.3234 (0.174) +
number of major classes	0.0416 (0.007) ***	0.0417 (0.007) ***	0.0494 (0.007) ***	0.0021 (0.009)	0.0023 (0.009)	0.0075 (0.009)
previous trials	-0.0002 (0)	-0.0002 (0)	-0.0003 (0)	-0.0006 (0) *	-0.0006 (0) *	-0.0005 (0) *
number of authors	0.0557 (0.003) ***	0.0556 (0.003) ***	0.0505 (0.003) ***	-0.0112 (0.004) **	-0.011 (0.004) **	-0.0115 (0.004) **
couplings already used by the firm	-0.0248 (0.013) +	-0.024 (0.013) +	-0.0267 (0.013) *	-0.0206 (0.017)	-0.0215 (0.017)	-0.0269 (0.017)
sub-classes already used by the firm	0.1409 (0.016) ***	0.1424 (0.016) ***	0.142 (0.016) ***	-0.1831 (0.022) ***	-0.1819 (0.022) ***	-0.1751 (0.022) ***
inventors by class	0.0654 (0.013) ***	0.0907 (0.019) ***	-0.0649 (0.015) ***	0.0005 (0) **	0.0007 (0) **	0.0004 (0) **
age	-0.0029 (0.005)	-0.0052 (0.005)	0.0156 (0.001) ***	0.0282 (0.007) ***	0.0281 (0.007) ***	-0.0145 (0.002) ***
R&D intensity	-0.4775 (0.233) *	-0.3867 (0.243)	-0.1383 (0.119)	-0.1961 (0.34)	-0.4232 (0.35)	-0.7069 (0.164) ***
Number of employees (log)	-0.0748 (0.018) ***	-0.0751 (0.018) ***	-0.0595 (0.009) ***	0.0616 (0.026) *	0.0539 (0.027) *	-0.056 (0.012) ***
Performance	0.0011 (0) **	0.0012 (0) **	0.0023 (0) ***	-0.0001 (0.001)	-0.0001 (0.001)	-0.0009 (0) *
1/N	-1.346 (0.202) ***	-1.349 (0.202) ***	-1.1813 (0.209) ***	0.7917 (0.267) **	0.7988 (0.267) **	0.9069 (0.265) **
1/N squared	1.1816 (0.299) ***	1.1845 (0.299) ***	1.0153 (0.308) **	-0.8948 (0.399) *	-0.9023 (0.399) *	-0.9833 (0.396) *
Interdependence (K)	-0.5918 (0.164) ***	-0.5591 (0.164) **	-1.0259 (0.256) ***	-0.413 (0.241) +	-0.4108 (0.243) +	-0.8422 (0.234) ***
K squared	0.1402 (0.033) ***	0.1387 (0.033) ***	0.6601 (0.331) *	-0.4008 (0.155) *	-0.42 (0.16) **	0.1091 (0.088)
K/N	0.1935 (0.175)	0.185 (0.175)	0.1452 (0.213)	0.7083 (0.254) **	0.7174 (0.255) **	0.6102 (0.25) *
skewness	2.1179 (0.5) ***	1.7111 (0.508) **	0.0062 (0.428)	-0.5641 (0.716)	-0.5515 (0.728)	0.0939 (0.568)
skewness squared	0.2903 (0.517)	0.018 (0.523)	-0.6418 (0.474)	-1.4034 (0.773) +	-1.2354 (0.782)	-0.9422 (0.655)
skewness x K	**-1.4171 (0.613)** *	**-1.2042 (0.606)** *	**-0.8199 (0.487)** +	-1.9797 (0.737) **	-1.9798 (0.74) **	-1.8654 (0.689) **
Log number of inventors - size	-0.2083 (0.021) ***	-0.2047 (0.032) ***	-0.0114 (0.026)	0.0524 (0.034)	0.0257 (0.043)	0.0441 (0.032)
skewness x size	-0.3062 (0.074) ***	-0.2486 (0.078) **	0.0163 (0.066)	0.0548 (0.107)	0.0564 (0.111)	-0.0013 (0.084)
KB's Clusterization		-0.2872 (0.155) +	0.4769 (0.127) ***		0.1178 (0.226)	-0.0053 (0.17)
KB's Density		0.6033 (1.085)	4.4257 (0.712) ***		5.5868 (1.394) ***	-0.2832 (0.746)
KB's blocks		0.0019 (0.001) +	-0.0021 (0.001) *		0.0004 (0.002)	-0.0005 (0.001)
SN's blocks		-0.0004 (0) **	0.0007 (0) ***		0 (0)	0.0002 (0) +
SN's density		-2.9359 (1.162) *	-1.3637 (0.832)		-5.1487 (1.62) **	-0.6906 (1.168)
SN's clusterization		0.3113 (0.145) *	0.3719 (0.093) ***		0.1195 (0.219)	-0.2118 (0.126) +
SN's mean tie		-0.0251 (0.017)	0.0175 (0.014)		-0.0431 (0.024) +	0.0067 (0.019)
Dummy Controls for main technological class / firm / granting year			No dummy for firm	Dummy for main technological class / firm / granting year		No dummy for firm
Num observations	89902	89902	89902			
Prob > chi2	0.0000	0.0000	0.0000			
Log pseudolikelihood	-289443.2	-289401.27	-291156.84			

(Standard Error) *** p<0.001 ** p<0.01 *p<0.05 +p<0.1

Appendix B

Appendix to Chapter 4

B.1 Measuring hierarchical decomposability

Matrix of distances

The starting point is the a firm's knowledge base represented by the coupling matrix, a symmetrical square matrix P_n where each element P_{ij} represents the coupling intensity of technological sub-classes i and j. Figure B.1 provides a simple example a of knowledge base composed by 10 patents classified in classes from A to F. This matrix is sorted in order to group around the main diagonal those elements with the strongest coupling values so that those values also progressively decay in further positions from the diagonal. The idea is to get as close as possible to a structure similar to Matrix 4.7. In order to do that, first I define a distance among the n elements of P_n. For that purpose, I build a matrix of direct distances among those elements with positive couplings DD_n such that:

$$DD_n = [dd_{ij}]_n \text{where} \begin{cases} dd_{ij} = max(P_n) + 1 - p_{ij} \text{ if } p_{ij} > 0 \\ dd_{ij} = 0 \text{ if } p_{ij} = 0 \end{cases} \tag{B.1}$$

DD_n defines direct distances whenever a pair of elements are coupled such that those elements with the highest coupling value have a direct distance of 1, while the minimum coupling value equals $max(P_n)$. Following the example, Figure B.1 shows this matrix in the second line. This transformation is done so the more intense the coupling between two elements is, the closer they are. Moreover, discrete values are used for the sake of computational simplicity. Once direct distances are defined, I proceed to calculate all distances in matrix D.

DD_n depicts a network where some elements are directly connected with lengths equal to dd_{ij}, while others are unconnected. In order to define distances among unconnected elements i and j, the shortest path (SP) between both of them is calculated among all the the possible indirect paths using direct connections with other nodes. I will define matrix D as:

$$D_n = [d_{ij}]_n \text{ where } d_{ij} \text{is the length of the SP between } i \text{ and } j \tag{B.2}$$

Figure B.1 depicts this matrix in the third line. This approach embraces the one of Yayavaram and Ahuja (2008). For example, they proposed that if two nodes i and j are loosely coupled but both of them are strongly coupled with a third node, they are integrated (i.e. close enough to be considered in a cluster). The matrix of distances D, by considering geodesics between each pair of elements, follows the same logic. It may happen that $dd_{ij} > d_{ij}$, or, in other words, that two nodes are closer than its direct distance, something that occurs when they are strongly connected with third parties.

If the network is composed by unconnected blocks of nodes (blocks A-B-C-D y E-F in Figure B.1), it is not possible to define a distance among those. Since all nodes of the network compose the knowledge base of a firm and consequently they are related, I assume the ties between unconnected blocks do not manifest since they are larger than the maximum distance among any other two. Hence, I consider them as:

$$D_n = [dd_{ij}]_n \text{ where } \begin{cases} dd_{ij} = \text{length of SP between } i \text{ and } j \\ \text{ if there is at least one path between them} \\ dd_{ij} = max(D_n) \text{ if there is not any path between } i \text{ and } j \end{cases} \tag{B.3}$$

Summing up, distances among technological sub-classes in the knowledge base are inversely related with their frequency of coupling in patents. Those that have never been together in a patent have a distance defined according to other sub-classes they associate with. Those sub-classes that are totally unconnected with others (directly and indirectly) are assumed to have the maximum defined distance on the rest of the knowledge base. Figure B.2 shows an example. Panel I plots a network of 25 elements. Panel II shows the matrix P associated with the network where instead of numbers, a colored scale shows with darker tones higher intensities and with lighter colors the opposite.

Coupling Matrix

$$P_{6x6} = \begin{pmatrix} 0 & 5 & 2 & 2 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 2 & 0 \end{pmatrix}$$

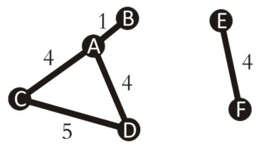
Patents

AB AB AB AD AC
ACD AB AB EF EF

Example Equation (8)

Direct distances

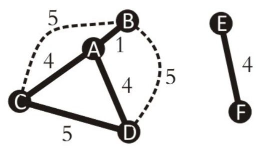
$$DD_{6x6} = \begin{pmatrix} 0 & 1 & 4 & 4 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 5 & 0 & 0 \\ 4 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 4 & 0 \end{pmatrix}$$



Example Equation (9)

Geodesic distances

$$D_{6x6} = \begin{pmatrix} 0 & 1 & 4 & 4 & 0 & 0 \\ 1 & 0 & 5 & 5 & 0 & 0 \\ 4 & 5 & 0 & 5 & 0 & 0 \\ 4 & 5 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 4 & 0 \end{pmatrix}$$



Example Equation (10)

Full matrix of distances

$$D_{6x6} = \begin{pmatrix} 0 & 1 & 4 & 4 & 5 & 5 \\ 1 & 0 & 5 & 5 & 5 & 5 \\ 4 & 5 & 0 & 5 & 5 & 5 \\ 4 & 5 & 5 & 0 & 5 & 5 \\ 5 & 5 & 5 & 5 & 0 & 4 \\ 5 & 5 & 5 & 5 & 4 & 0 \end{pmatrix}$$

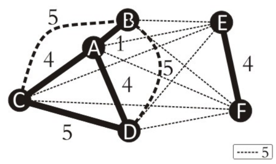


Figure B.1: Example to Equations B.1, B.2 and B.3.

Hierarchical clustering

Once defined the matrix of distances D_n from Equation B.1, its n elements are agglomeratively clustered by minimizing the average distance within groups (equivalent of maximizing average coupling intensity within groups). The underlying logic of this process is to calculate how elements group as we relax the distance. As groups include more elements, hierarchies are revealed as it can be seen at Panel I from Figure B.2. Panel V shows the resulting clustering tree of elements from P_n. In this example, the whole network is divided into two groups that are divided into 4 groups, and so on. Panel III of Figure B.2 shows matrix P_n as showed in Panel II but sorted according to the sequence of elements after clustering as they appear at horizontal axis in Panel V. This new matrix PS_n groups along the main diagonal the most intense couplings values so they decrease as we get further.

Measurement

Once defined the grouping sequence, I can proceed to calculate a measurement for capturing the decomposability of a group of interrelated elements. Assume S is a system composed by $n(S)$[1] elements and we need to analyze all of them in order to find a specific piece $x \in S$, as if S was a library and x a specific book. Assume also that we do not know which piece is x but we can figure it out after analyzing all elements. We should analyze all $n(S)$ elements. Now suppose that we divide the entire set by some criterion that maximize internal homogeneity and external heterogeneity among G groups g_i's. Now, instead of analyzing $n(S)$ elements, analyzing the G groups would be enough to know where element x is, and then, we will proceed to analyze the $n(g_i)$ elements withing g_i. Hence, the effort of analyzing the whole set would be $e = G + n(gi) < n(S)$. If the set was a library, the effort of finding certain business book could be considerably reduced from the effort of analyzing the entire library by dividing it in broad categories so we can easily ignore entire sections of psychology, cooking or literature books, for example.

With this intuition in mind, I call e the effort of analyzing a set S of $n(S)$ interrelated elements. When the system is zero decomposable, the effort of analyzing S is equal to $n(S)$. However, if S is decomposable in some degree, S could be divided in G internally homogeneous but externally heterogeneous groups. If P is the matrix that describe the intensity of relations among the $n(S)$ elements, it can be stated as Equation B.4,

[1] From now on I will use the cardinal of the set S -number of elements- as a function $n(S)$ for the sake of notation simplicity

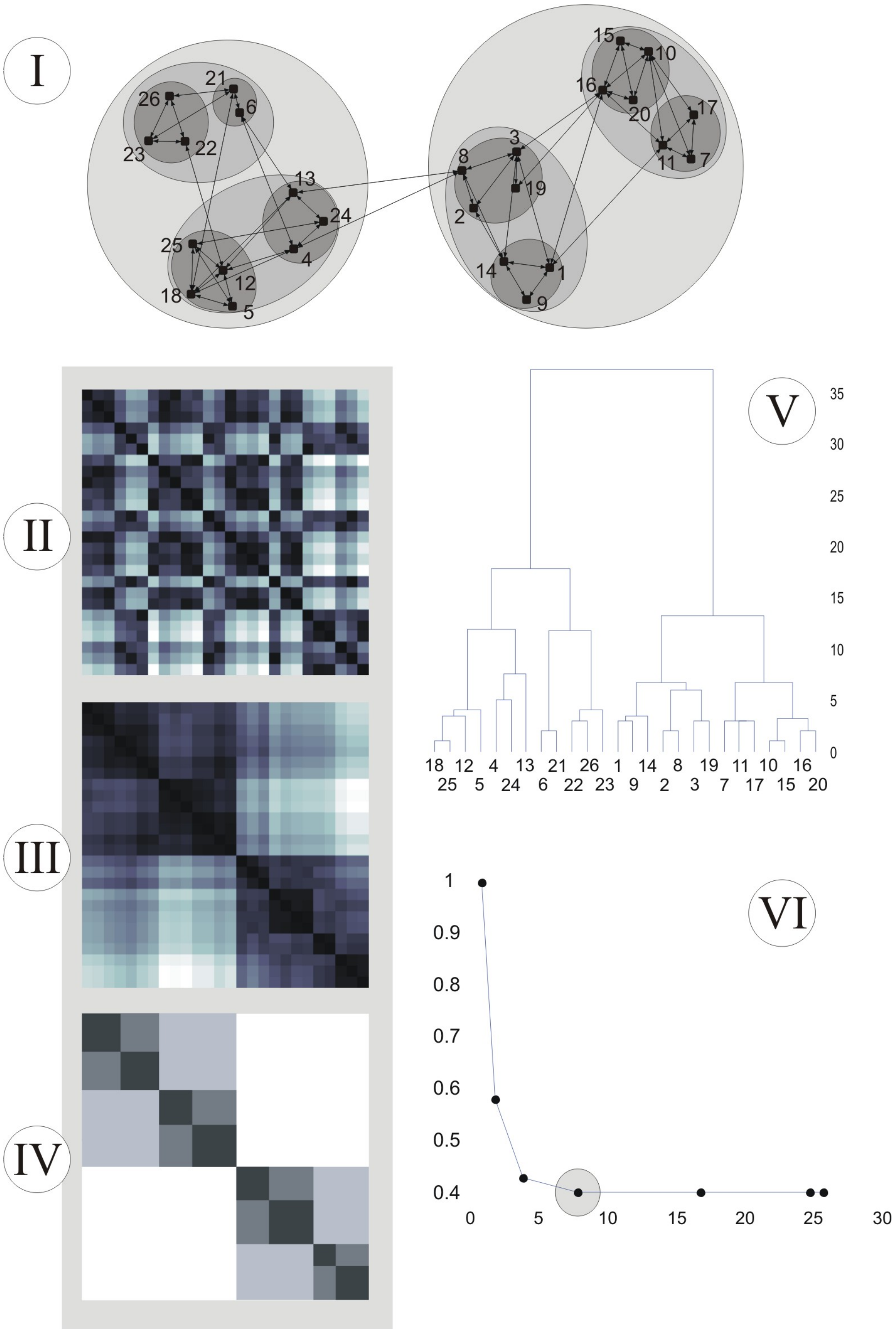


Figure B.2: Example of Hierarchical Near Decomposability where $G < n(S)$, $P_{i=j}$ are the intensities withing groups (as average), and $P_{i\neq j}$ are the intensities among groups.

$$P_{GxG} = \begin{pmatrix} P_{11} & \cdots & P_{1j} & \cdots & P_{1G} \\ \vdots & \ddots & \vdots & & \vdots \\ P_{i1} & \cdots & P_{ij} & \cdots & P_{iG} \\ \vdots & & \vdots & \ddots & \vdots \\ P_{G1} & \cdots & P_{Gj} & \cdots & P_{GG} \end{pmatrix} \tag{B.4}$$

With this structure I define the degree of decomposability of S as:

$$d(S) = \begin{cases} 1 - \dfrac{max(P_{i\neq j})}{min(P_{i=j})} \text{ if } max(P_{i\neq j}) \leqslant min(P_{i=j}) \\ 0 \text{ if } max(P_{i\neq j}) > min(P_{i=j}) \end{cases} \tag{B.5}$$

In words, I look for the minimum level of intensities of inter-group relations and I compare it with the maximum level of intensities of intra-group relations. If this coefficient is larger than 1, it means that similarities withing groups are weaker than among groups, i.e. there is not a clear distinction among groups. I take maximum and minimum values instead of considering averages in order to penalize *any* inter-group relation too much strong. If $d = 0$, then S is zero decomposable since there is not any gain in homogeinity by grouping elements. On the contrary, if $d = 1$, then, the system is fully decomposable since groups are not related *among* them, but all relations are *within* them.

As said before, if S is zero decomposable ($d = 0$), then, the effort of analyzing the group would be n. However, if it is decomposable in some degree ($d > 1$), the effort might be smaller. For analyzing certain element $s_i \in S$, the effort could be described as:

$$e(s_i) = (1 - d_{(S)})n_{(S)} + d_{(S)}(G_{(S)} + n_{(S_g)}) \text{ such that } s_i \in S_g \; g = 1, ..., G \tag{B.6}$$

This equation weights the two possible scenarios. With a weight of $(1 - d)$, it contemplates the scenario of analyzing the whole group of $n(S)$ elements. The less decomposable S, the larger the possibility of analyzing the full set of $n(S)$ elements. The other possible situation is weighted with d. If S is fully decomposable, then, we will analyze only the number of groups which S is divided in (G), plus the number of elements of the group which s_i belongs to.

Since the effort of analyzing each element s_i varies depending on which sub-group S_g contains it, I will define the effort of analyzing S as the average effort of each element.

$$e(S) = \frac{1}{n_{(S)}} \sum_{i=1}^{n_{(S)}} e(s_i) = (1 - d_{(S)})n_{(S)} + d_{(S)} \left(G_{(S)} + \sum_{g=1}^{G_{(S)}} n_{(S_g)} \frac{n_{(S_g)}}{n_{(S)}} \right) \tag{B.7}$$

As equation B.7 shows, the effort of analyzing S is equal to its number of elements $n(S)$ if it is zero decomposable, or it is equal to the number of groups plus the weighted average of the number of elements of each of the sub-groups of S in which it decomposes.

But what happens if S decomposes in G sub-groups S_g's and some of these sub-groups *also decompose* into sub-groups. If sub-group S_g has some degree of decomposability, then, the effort of analyzing S_g will not be $n(S_g)$ but $e(S_g)$. Thus, in case of hierarchical clustering, equation B.7 can be written as:

$$e(S, L) = \frac{1}{n_{(S)}} \sum_{i=1}^{n_{(S)}} e(s_i) = (1 - d_{(S)})n_{(S)} + d_{(S)} \left(G_{(S)} + \sum_{g=1}^{G_{(S)}} e(S_g, L-1) \frac{n_{(S_g)}}{n_{(S)}} \right) \tag{B.8}$$

For each sub-group S_g of S, the same logic applies since it can be decomposed into smaller sub-groups. I include the parameter L as indicator of how many times we divide the system S. If S is large and heteregenous enough, we might be able to divide the groups many times until we arrive to sub-groups with $d = 0$.

Finally, I define the relative effort of analyzing S as $e(S, L)$ divided over n. The value of $E(S, L)$ would indicate which percentage of the size of S remains to analyze after decomposing the system.

$$E(S, L) = \frac{e(S, L)}{n} \tag{B.9}$$

If S is zero decomposable, $E(S, L)$ is 1. In other words, there is not gain of dividing the system into sub-systems. The smaller the value of E, the larger the gains of decomposing S. In the example portrayed in Figure B.2, $E(S, L)$ is showed in Panel VI as sequentially adding more levels of sub-division. As it can be observed, when the group is not divided at all, $E(S,1) = 1$. When it is divided into 2 sub-systems, the effort of analyzing the group is reduced to $E(S,2) = 58\%$. When both sub-systems are divided into two sub-systems each, the effort goes down to $E(S,3) = 42\%$, and so on. As I proceed adding levels of clustering, there might be a lost of efficiency. Then, I look for the optimum number of levels I should cluster S, L^* such that $E(S,L^*) = min(E(S,L^*))$. Panel VI shows the minimum level of effort at the third division in that particular example. Panel IV shows the matrix of intensities P (at Panel II and sorted at Panel III), arranged and clustered according to the optimum level of hierarchical clusterization. The system of 25 elements now can be easily observed, comprehend and analyzed as a system composed by 8 sub-systems, grouped into 4 and then into 2. Panel IV would be the intuitive interpretation of Panel III by a human mind.

$E(S,L^*)$ is a measurement that tries to capture the gains of simplifying throughout hierarchical clustering systems with interrelated components. It is independent of the size of the system, and it includes much valuable information about its structure such the size of the group and each sub-group in each division, differences in intensities of intra and inter relations among groups in each level and group, and the number of levels we decompose the structurre. $E(S,L)$ can be larger than 1 meaning that decomposing S into L levels can be so adverse that demand more effort than not decomposing. However, $E^*(S,L)$ is always equal or smaller than 1 since in the worst scenario we can always choose $E(S,L=0) = 1$.

Figure B.3 shows some examples of hierarchical clusterization and the E index measurement. The first line of squares shows the similarity matrix sorted to maximize similarities around the main diagonal as explained before. Darker tones indicate very similar elements, while lighter ones the opposite. The second line shows the intensity matrices after being clustered at the optimum level. Here, dark tones indicate intense couplings while light ones, the opposite. The third line pictures the agglomerative cluster-tree. Finally, between the two lines of matrices, there is E index for each case.

On the first column there is a case of absolute impossibility of clustering. This structure does not have any group of element that increase the internal homogeneity by increasing the external heterogeneity. As result, the minimum value of E is achieved when its components are not clustered at all: it is 1 and it means that there is not gain in clustering. At column II, a network with links randomly formed is shown. As it can be observed there is gain of clustering but quite small. The second line of matrices shows the optimum clusterization and for this example the internal intensity of groups and

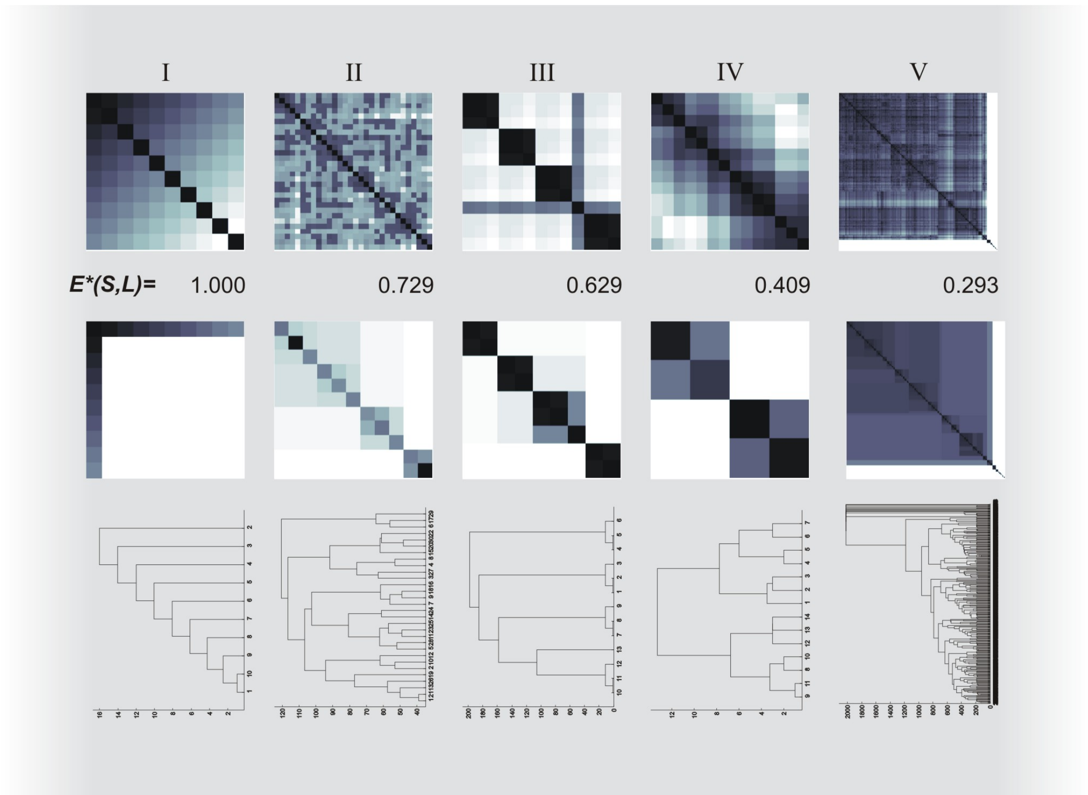


Figure B.3: Examples of HD

its homogeneity in size is low. Column III shows a network strongly clustered at one level while Column IV shows a network of similar size clustered at two levels. The effort of analyzing IV is inferior than analyzing III. Finally, Column V shows the coupling network of Intel in 1993-1995, a network composed by 312 elements (technological subclasses that the firm has coupled by patents produced). The seemingly quite complex structure maximally simplifies when clustered as the second grey-scaled matrix shows. The gains of HND in this late example are larger than in the previous due to the size of the matrix.

Finally, in order to have a measurement that increases as it captures the concept involved, the hierarchical near-decomposability of a knowledge base will be measured as:

$$HD(S,L) = 1 - E(S,L) \tag{B.10}$$

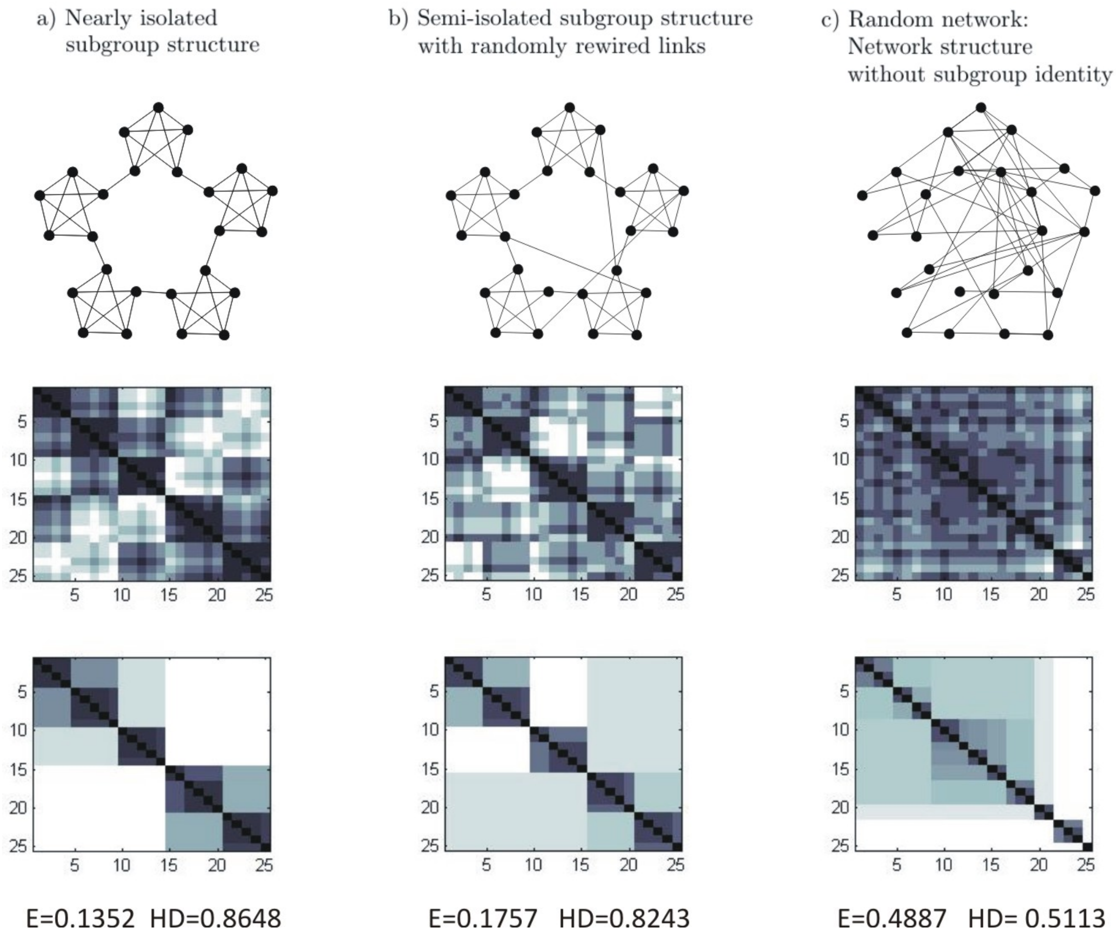


Figure B.4: Example used by Fang et al. (2010) based on variations on Watts (1999)'s Connected Caveman Model. The structure with isolated subgroups shows a higher hierachical decomposability than the semi-isolated, and much more than the random network.

B.2 Decomposability index by Yayavaram and Ahuja (2008).

Yayavaram and Ahuja (2008) propose using a modified version of a standard cluster coefficient for networks. While the standard measurement averages the density of the sub-networks associated to each node (defined as those nodes connected to the focal one), the decomposability index proceeds the same way but considering the density of strong ties in sub-networks. This differentiation focuses on knowledge combination that the focal firm prioritizes above the common beliefs within the sector. This distinction between weak and strong ties is defined in the following paragraphs.

For each firm, its knowledge base at time t is determined as the network of classes where each couple of nodes is coupled by the number of firm's patents that has been produced withing the 3 previous years to t and are classified simultaneously in those two nodes-classes. Then, median values of non-zero coupling are calculated for each firm-period. Following Yayavaram and Ahuja (2008), in order to isolate median values from size and time effect, the logarithm of median values of positive couplings are regressed against year and the logarithm of the quantity of patents of knowledge bases. Results are shown at Table B.2.

Dependent variable: log (median(couplings))

constant	-14.7780 (.0004)
year/month	0.0076 (.0016)
log # patents	0.0223 (.7660)

All significant with p-value < 0.005

Secondly, Yayavaram and Ahuja (2008) distinguish between *within the cluster ties* (WCT) and *out of the cluster* (OCT) ties. A tie is considered within the cluster whenever: i) it is strong tie, and among the nodes it links there is a least one common node to which both nodes are connected with; ii) it is a strong tie and both connected nodes do not have any other connection to any other node; iii) it is a weak tie but both connected nodes are strongly connected to a common node. These three cases are plotted in Figure B.5.

Thirdly, after classifying all ties as within or out the cluster, it is calculated the *integration* of each node. In order to do that, for the sub-network of each node including all

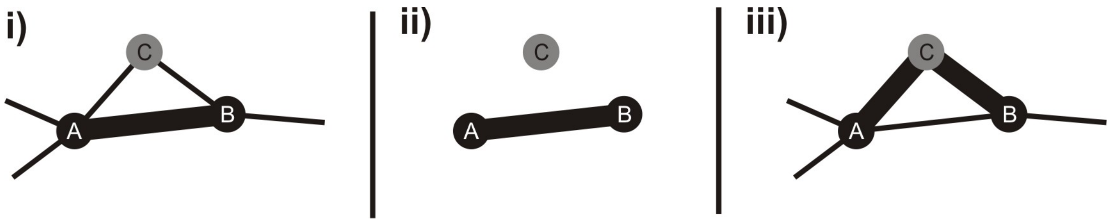


Figure B.5: Criterion for distinguishing within-the-cluster ties. A and B are considered connected by a within-cluster tie in each of these 3 cases.

links between the focal node and its neighbors as well as ties between neighbors, equation B.11 is calculated. Nodes without any ties are considered to have zero integration.

$$Integration_{node} = \frac{\text{OCT}}{\text{possible number of ties}} \tag{B.11}$$

Finally, the decomposability of the knowledge base network is calculated as (1-weighted sum of the integration of each node), using as weights the percentage of patents that belong to each node. In other words, the more integrated are the nodes of the network in average, the less decomposable the network is.

This methodology rises certain issues. On the one hand, nodes without any ties are considered with zero integration. This means that certain node which is not connected to the network it is not integrated. However, on the other hand, a node whose sub-network is composed, for example, by all strong ties and with full density, according to the previous Equation, it is zero integrated as well. This is a clear contradiction. A node with no ties at all cannot be as integrated as a node that belongs to a network fully connected by strong ties. The Equation B.11 should corrected to:

$$Integration_{node} = \frac{\text{WCT}}{\text{possible number of ties}} \tag{B.12}$$

This way, the more *within the cluster* ties among the sub-network of certain node, the more integrated is considered the focal node. As expected, the alternative version is negative correlated with the original one. In this research, Yayavaram and Ahuja (2008)'s decomposability index is considered as Equation B.12.

Finally, this paper also changes the 'scale' of the measurement. Instead of analyzing how firms couple technological classes, the focus is on technological aggregated sub-classes. While Yayavaram and Ahuja (2008) use 30 classes (i.e. the maximum number of node for a network is 30), I consider 100 classes divided into 1,871 aggregated sub-classes. I work at this level of aggregation in order to expand the too-much-broad categories of classes but not so much as to get weak coupling between sub-classes. This approach allows analyzing much more fine-grained knowledge structures with more demanding measurements.

B.3 Descriptive Statistics

Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max
citations	10.3903	15.9623	0	314
prior citations	11.8633	18.7141	0	568
number of authors	2.2514	1.4511	0	26
number of major classes	1.4518	.7200	1	10
previous trials	6.0053	30.1764	0	889
couplings already used by the firm	.3664	.4188	0	1
sub-classes already used by the firm	.6524	.3849	0	1
single sub-class	3.660	2.3389	1	35
number of sub-classes (N)	0.1019	0.3025	0	1

dummies main class	100 classes
dummies publication year	1981-2006
Number of observations	110,693

Correlation Matrix

Variable	citations	prior art	authors	Nº classes	repeated trials	prev. combinations	prev. piece	Nº sub-classes
citations	1.0000							
prior citations	0.1155*	1.0000						
number of authors	0.0635*	0.0561*	1.0000					
number of major classes	0.0524*	0.0112*	-0.0058*	1.0000				
previous trials	-0.0340*	-0.0099*	0.0025*	-0.1161*	1.0000			
couplings already used by the firm	-0.0153*	0.0662*	0.0294*	-0.1989*	0.2016*	1.0000		
sub-classes already used by the firm	0.0158*	0.0685*	0.0528*	-0.1014*	0.0902*	0.5448*	1.0000	
number of sub-classes (N)	0.0626*	0.0426*	0.0156*	0.4677*	-0.1859*	-0.2010*	-0.0183*	1.0000

*p<0.05

Figure B.6: Descriptive statistics for Table 4.1

Descriptive statistics

Variable	Mean	Std. Dev.	Min	Max	VIF
iq	0.1065	0.0643	0.012	0.436	
patents	88.7019	207.4707	5.000	1863.000	2.98
R&D intensity	0.1184	0.1273	0.003	2.589	3.26
Log of employes	1.0848	1.7307	-3.381	4.604	2.42
Return on assets	-2.9340	49.2843	-1393.712	158.680	2.68
KB's number of sub-classes	91.6238	112.8645	10.000	687.000	5.30
Inventors by components	1.4635	0.7912	0.345	5.472	2.82
Years patenting	10.2700	5.9341	0.000	24.000	1.26
ND of KB	0.5260	0.1430	0.000	0.786	2.18
ND of SN	0.5056	0.1958	0.000	0.861	2.47
K of KB	0.1759	0.0325	0.111	0.268	1.17
Size SN	192.4766	386.3196	10.000	3651	

Number of groups	137
Number of observations	961

Correlation Matrix

	iq	patents	R D intensity	Log of employes	Return on assets	KB's number of sub-classes	Inventors by components	Years patenting	ND of KB	ND of SN	K of KB	Size SN
iq	1,000											
patents	0,605	1,000										
R&D intensity	-0,037	-0,069	1,000									
Log of employes	0,395	0,408	-0,372	1,000								
Return on assets	0,107	0,051	-0,735	0,198	1,000							
KB's number of sub-classes	0,616	0,795	-0,058	0,588	0,061	1,000						
Inventors by components	0,403	0,467	-0,173	0,573	0,083	0,623	1,000					
Years patenting	0,207	0,179	-0,178	0,185	0,061	0,297	0,380	1,000				
ND of KB	0,412	0,307	-0,041	0,454	0,082	0,523	0,233	0,198	1,000			
ND of SN	0,436	0,338	-0,103	0,482	0,072	0,529	0,634	0,380	0,585	1,000		
K of KB	0,068	-0,173	0,125	-0,243	0,047	-0,247	-0,106	-0,027	-0,174	-0,135	1,000	
Size SN	0,510	0,725	-0,058	0,520	0,055	0,911	0,732	0,304	0,368	0,447	-0,181	1,000

Figure B.7: Descriptive statistics for Table 4.2

B.4 Extended Results: summary statistics

	Mean	Std Dev	Min	Max
cit	9,915	15,948	0	313
prior art	11,972	18,030	0	378
dummy single class	0,102	0,303	0	1
Number of main classes	1,443	0,711	1	8
Repeated trials control	6,591	33,877	0	889
N	3,642	2,345	1	35
K	0,220	0,086	0,067	2,625
number of authors	2,245	1,437	1	26
previous combinations used	0,429	0,419	0	1
subclasses used	0,778	0,319	0	1
number of subclasses	420,963	138,626	134	701
inventors by class	2,661	1,173	0,930	5,684
age	13,839	4,740	2	24
R&D intensity	0,095	0,047	0,014	0,324
Number of employees (log)	3,152	0,865	0,919	4,604
Performance	3,457	18,297	-100,922	158,680
Decomposability	0,652	0,069	0,442	0,776
Social Decomp	0,697	0,121	0,153	0,861
SN's size	1201,345	833,372	177	3651
KB's Clusterization	0,563	0,034	0,432	0,692
KB's Density	0,020	0,005	0,012	0,044
KB's blocks	21,311	6,760	4	46
SN's blocks	222,834	160,671	15	686
SN's density	0,006	0,005	0,001	0,029
SN's clusterization	0,517	0,054	0,366	0,708
SN's mean tie	1,917	0,586	1,158	3,715
control class		100 classes		
control firm		20 firms		
control year		22 years		

Figure B.8

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	cit	1.0000								
(2)	prior art	0.0893	1.0000							
(3)	Number of main classes	0.0452	0.0199	1.0000						
(4)	Repeated trials control	-0.0368	-0.0202	-0.1140	1.0000					
(5)	N	0.0537	0.0552	0.4314	-0.1838	1.0000				
(6)	K	-0.0220	-0.0763	-0.0537	0.1366	-0.2959	1.0000			
(7)	number of authors	0.0406	0.0192	-0.0084	-0.0111	0.0065	-0.0914	1.0000		
(8)	previous combinations used	-0.0194	0.0711	-0.2337	0.2040	-0.2132	-0.0102	0.0117	1.0000	
(9)	subclasses used	0.0224	0.0785	-0.0962	0.0760	-0.0081	-0.1578	0.0386	0.4714	1.0000
(10)	number of subclasses	0.0133	0.1345	0.0265	0.0066	0.0138	-0.0301	-0.0390	0.0779	0.1248
(11)	inventors by class	-0.0334	-0.0905	0.0414	0.0355	-0.0390	0.1035	0.0740	-0.0189	0.0452
(12)	age	-0.0079	0.0124	0.0227	0.0347	-0.0041	-0.0222	0.0585	0.0484	0.0968
(13)	R&D intensity	-0.0359	-0.0718	0.0208	-0.0149	-0.0454	0.1445	0.0044	-0.1059	-0.1353
(14)	Number of employees (log)	0.0086	-0.0514	0.0661	-0.0091	-0.0342	0.1487	0.0140	-0.0537	-0.0409
(15)	Performance	0.0525	-0.0647	0.0049	-0.0049	-0.0201	0.0474	0.0023	-0.0004	0.0195
(16)	Decomposability	0.0378	-0.0910	0.0140	-0.0174	-0.0308	0.1051	-0.0157	-0.1120	-0.1145
(17)	Social Decomp	0.0305	0.0006	0.0465	-0.0139	-0.0326	0.1270	-0.0215	-0.0412	-0.0251
(18)	Ahujas's Decomp	-0.1168	-0.1979	0.0058	0.0093	-0.0608	0.1678	0.0379	-0.1401	-0.1976
(19)	SN's size	-0.0212	-0.0197	0.0394	0.0309	-0.0212	0.0544	0.0358	0.0146	0.0778
(20)	KB's Clusterization	0.0524	0.1807	0.0113	-0.0052	0.0346	-0.0507	-0.0183	0.1094	0.1116
(21)	KB's Density	0.0162	0.0151	-0.0055	0.0161	0.0135	-0.0313	0.0476	0.0515	0.0592
(22)	KB's blocks	-0.0525	-0.0948	0.0122	-0.0029	-0.0183	0.0414	-0.0035	-0.0683	-0.0811
(23)	SN's blocks	-0.0193	-0.0399	0.0494	0.0161	-0.0384	0.1256	0.0038	-0.0341	0.0043
(24)	SN's density	-0.0091	-0.0201	-0.0546	0.0128	0.0351	-0.1417	0.0213	0.0485	0.0270
(25)	SN's clusterization	-0.0396	-0.1919	0.0248	0.0247	-0.0226	0.0512	0.1558	-0.0738	-0.0561
(26)	SN's mean tie	0.0228	0.2570	-0.0388	0.0013	0.0653	-0.2099	-0.0563	0.1832	0.2090

		(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
(10)	number of subclasses	1.0000								
(11)	inventors by class	0.4868	1.0000							
(12)	age	0.3236	0.5091	1.0000						
(13)	R&D intensity	0.0176	0.0807	0.3024	1.0000					
(14)	Number of employees (log)	0.5993	0.6779	0.0897	-0.0741	1.0000				
(15)	Performance	-0.0006	0.1560	-0.1077	-0.4936	0.2966	1.0000			
(16)	Decomposability	0.0950	0.1457	-0.0140	0.2611	0.1934	-0.0056	1.0000		
(17)	Social Decomp	0.4361	0.3664	0.3941	0.4858	0.3281	-0.2239	0.3594	1.0000	
(18)	Ahuja's Decomp	-0.2484	0.1425	-0.0344	0.4123	0.0542	-0.1084	0.2603	0.1406	1.0000
(19)	SN's size	0.7544	0.9191	0.5083	0.0624	0.7106	0.1308	0.1643	0.3982	-0.0006
(20)	KB's Clusterization	-0.1057	-0.3439	-0.0992	-0.2016	-0.2360	-0.0364	-0.2931	-0.1116	-0.3923
(21)	KB's Density	-0.6827	-0.2013	-0.0350	-0.1866	-0.5063	-0.0054	-0.4149	-0.4107	-0.1687
(22)	KB's blocks	0.3012	0.2623	0.0513	0.0692	0.3622	0.0180	0.1478	0.1167	0.2474
(23)	SN's blocks	0.7367	0.8629	0.3889	0.1368	0.7753	0.1364	0.2449	0.4311	0.1247
(24)	SN's density	-0.7026	-0.5936	-0.3544	-0.3713	-0.6909	0.0974	-0.3452	-0.7415	-0.1308
(25)	SN's clusterization	-0.0768	0.4464	0.3306	0.2047	0.2134	-0.1219	0.0866	0.3088	0.3008
(26)	SN's mean tie	0.2431	-0.3995	0.0675	-0.1105	-0.3613	-0.1585	-0.1827	0.0064	-0.5017

		(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
(19)	SN's size	1.0000							
(20)	KB's Clusterization	-0.3342	1.0000						
(21)	KB's Density	-0.4352	0.4734	1.0000					
(22)	KB's blocks	0.3279	-0.6015	-0.5360	1.0000				
(23)	SN's blocks	0.9549	-0.4103	-0.5422	0.3826	1.0000			
(24)	SN's density	-0.6586	0.2446	0.6411	-0.3296	-0.7274	1.0000		
(25)	SN's clusterization	0.2511	-0.1367	0.0840	0.0009	0.1614	-0.1723	1.0000	
(26)	SN's mean tie	-0.2128	0.4454	0.0392	-0.3124	-0.3234	0.0935	-0.4532	1.0000

Figure B.9

References

Acemoglu, D., Ozdaglar, A., and ParandehGheibi, A. (2010). Spread of (mis) information in social networks. *Games and Economic Behavior*, 70(2):194–227.

Adler, P. S. and Kwon, S.-W. (2002). Social capital: Prospects for a new concept. *Academy of management review*, 27(1):17–40.

Allen, T. J. (1984a). Managing the flow of technology: Technology transfer and the dissemination of technological information within the r&d organization. *MIT Press Books*, 1.

Allen, T. J. (1984b). Managing the flow of technology: Technology transfer and the dissemination of technological information within the r&d organization. *MIT Press Books*, 1.

Almeida, P. and Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management science*, 45(7):905–917.

Arthur, W. B. (1994). Inductive reasoning and bounded rationality. *The American economic review*, pages 406–411.

Arthur, W. B. (2009). *The nature of technology: What it is and how it evolves*. Simon and Schuster.

Back, T., Fogel, D. B., and Michalewicz, Z. (1997). *Handbook of evolutionary computation*. IOP Publishing Ltd.

Baldwin, C. Y. and Clark, K. B. (2000). *Design rules: The power of modularity*, volume 1. Mit Press.

Bell, G. G. (2005). Clusters, networks, and firm innovativeness. *Strategic management journal*, 26(3):287–295.

Berlow, E. L., Dunne, J. A., Martinez, N. D., Stark, P. B., Williams, R. J., and Brose, U. (2009). Simple prediction of interaction strengths in complex food webs. *Proceedings of the National Academy of Sciences*, 106(1):187–191.

Birdsell, J. A. and Wills, C. (2003). The evolutionary origin and maintenance of sexual recombination: a review of contemporary models. In *Evolutionary Biology*, pages 27–138. Springer.

Bolton, P. and Dewatripont, M. (1994). The firm as a communication network. *The Quarterly Journal of Economics*, pages 809–839.

Bouty, I. (2000). Interpersonal and interaction influences on informal resource exchanges between r&d researchers across organizational boundaries. *Academy of Management Journal*, 43(1):50–65.

Bower, G. H. and Hilgard, E. R. (1981). Theories of learning.

Burt, R. (1992). *Structural Holes*. Havard University Press, Cambridge MA.

Caldwell, B. S. and Wang, E. (2010). Delays and user performance in human-computer-network interaction tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*.

Candiani, J. A. (2012). The clash of social and knowledge spaces: The assumed ismorphism under the hood.

Carley, K. M. (1990). Coordinating the success: trading information redundancy for task simplicity. In *System Sciences, 1990., Proceedings of the Twenty-Third Annual Hawaii International Conference on*, volume 4, pages 261–270. IEEE.

Chen, B. L., Hall, D. H., and Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12):4723–4728.

Chen, Z. and Guan, J. (2010). The impact of small world on innovation: An empirical study of 16 countries. *Journal of Informetrics*, 4(1):97–106.

Cohen, W. and Levinthal, D. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative science quarterly*, pages 128–152.

Cowan, R. and Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of economic Dynamics and Control*, 28(8):1557–1575.

Cowan, R. and Jonard, N. (2009). Knowledge portfolios and the organization of innovation networks. *Academy of Management Review*, 34(2):320–342.

Cowan, R., Jonard, N., and Özman, M. (2004). Knowledge dynamics in a network industry. *Technological Forecasting and Social Change*, 71(5):469–484.

Cross, R., Parker, A., Prusak, L., and Borgatti, S. P. (2001). Knowing what we know: Supporting knowledge creation and sharing in social networks. *Organizational dynamics*, 30(2):100–120.

Csaszar, F. A. (2013). An efficient frontier in organization design: Organizational structure as a determinant of exploration and exploitation. *Organization Science*, 24(4):1083–1101.

Csikszentmihalyi, M. (1997). Flow and the psychology of discovery and invention. *HarperPerennial, New York*.

Damanpour, F. (1996). Organizational complexity and innovation: developing and testing multiple contingency models. *Management science*, 42(5):693–716.

DeCanio, S. J., Dibble, C., and Amir-Atefi, K. (2000). The importance of organizational structure for the adoption of innovations. *Management science*, 46(10):1285–1299.

DeCanio, S. J. and Watkins, W. E. (1998). Information processing and organizational structure. *Journal of economic behavior & organization*, 36(3):275–294.

Dunbar, R. I. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, 22(6):469–493.

Einstein, A. (1934). On the method of theoretical physics. *Philosophy of science*, 1(2):163–169.

Eisenberg, J. F. (1981). The mammalian radiation: An analysis of trends in evolution, adaptation, and behavior. *Uni-versity of Chicago Press, Chicago*.

Ethiraj, S. K. and Levinthal, D. (2004). Modularity and innovation in complex systems. *Management Science*, 50(2):159–173.

Fang, C., Lee, J., and Schilling, M. A. (2010). Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science*, 21(3):625–642.

Feld, S. L. (1981). The focused organization of social ties. *American journal of sociology*, pages 1015–1035.

Felsenthal, D. S. (1980). Applying the redundancy concept to administrative organizations. *Public administration review*, pages 247–252.

Ferligoj, A. and Hlebec, V. (1999). Evaluation of social network measurement instruments. *Social networks*, 21(2):111–130.

Fleming, L. and Frenken, K. (2007). The evolution of inventor networks in the silicon valley and boston regions. *Advances in Complex Systems*, 10(01):53–71.

Fleming, L., King, C., and Juda, A. I. (2006). Small worlds and regional innovation. *Available at SSRN 892871*.

Fleming, L. and Sorenson, O. (2001). Technology as a complex adaptive system: evidence from patent data. *Research Policy*, 30(7):1019–1039.

Fleming, L. and Sorenson, O. (2004). Science as a map in technological search. *Strategic Management Journal*, 25(8-9):909–928.

Fleming, L. and Waguespack, D. M. (2007). Brokerage, boundary spanning, and leadership in open innovation communities. *Organization Science*, 18(2):165–180.

Ford, R. C. and Randolph, W. A. (1992). Cross-functional structures: A review and integration of matrix organization and project management. *Journal of management*, 18(2):267–294.

Gabbay, S. M. and Zuckerman, E. W. (1998). Social capital and opportunity in corporate r&d: The contingent effect of contact density on mobility expectations. *Social Science Research*, 27(2):189–217.

Galbraith, J. R. (1977). Organization design: An information processing view. *Organizational Effectiveness Center and School*, 21.

Gloor, P., Laubacher, R., Dynes, S., and Zhao, Y. (2003). Visualization of communication patterns in collaborative innovation networks-analysis of some w3c working groups. In *Proceedings of the twelfth international conference on Information and knowledge management*, pages 56–60. ACM.

Glynn, M. A. (1996). Innovative genius: A framework for relating individual and organizational intelligences to innovation. *Academy of management review*, 21(4):1081–1111.

Gonçalves, B., Perra, N., and Vespignani, A. (2011). Modeling users' activity on twitter networks: Validation of dunbar's number. *PloS one*, 6(8):e22656.

Granovetter, M. (1973). The strength of weak ties. *American journal of sociology*, pages 1360–1380.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological theory*, 1(1):201–233.

Granovetter, M. (1985). Economic action and social structure: the problem of embeddedness. *American journal of sociology*, pages 481–510.

Granovetter, M. (1992). Problems of explanation in economic sociology. *Networks and organizations: Structure, form, and action*, 25:56.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic management journal*, 17:109–122.

Grippa, F., Zilli, A., Laubacher, R., and Gloor, P. (2006). E-mail may not reflect the social network. In *Proceedings of the North American Association for Computational Social and Organizational Science Conference*.

Guimera, R., Danon, L., Diaz-Guilera, A., Giralt, F., and Arenas, A. (2003). Self-similar community structure in a network of human interactions. *Physical Review E*, 68(6):065103.

Guimera, R., Danon, L., Diaz-Guilera, A., Giralt, F., and Arenas, A. (2006a). The real communication network behind the formal chart: Community structure in organizations. *Journal of Economic Behavior & Organization*, 61(4):653–667.

Guimera, R., Danon, L., Diaz-Guilera, A., Giralt, F., and Arenas, A. (2006b). The real communication network behind the formal chart: Community structure in organizations. *Journal of Economic Behavior & Organization*, 61(4):653–667.

Guimera, R., Uzzi, B., Spiro, J., and Amaral, L. A. N. (2005). Team assembly mechanisms determine collaboration network structure and team performance. *Science*, 308(5722):697–702.

Guler, I. and Nerkar, A. (2012). The impact of global and local cohesion on innovation in the pharmaceutical industry. *Strategic Management Journal*, 33(5):535–549.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The nber patent citation data file: Lessons, insights and methodological tools. Technical report, National Bureau of Economic Research.

Hall, B. H. and Ziedonis, R. H. (2001). The patent paradox revisited: an empirical study of patenting in the us semiconductor industry, 1979-1995. *RAND Journal of Economics*, pages 101–128.

Hansen, M. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative science quarterly*, 44(1):82–111.

Hansen, M. T. and Løvås, B. (2004). How do multinational companies leverage technological competencies? moving from single to interdependent explanations. *Strategic Management Journal*, 25(8-9):801–822.

Hargadon, A. and Sutton, R. I. (1997). Technology brokering and innovation in a product development firm. *Administrative science quarterly*, pages 716–749.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society*, pages 1251–1271.

Hausmann, R., Hidalgo, C., Bustos, S., Coscia, M., Chung, S., Jimenez, J., Simoes, A., and Yildirim, M. (2011). The atlas of economic complexity: Mapping paths to prosperity. *Center for International Development at Harvard University and Macro Connections IT Media Lab, http://atlas. media. mit. edu*.

Hayek, F. A. (1945). The use of knowledge in society. *The American economic review*, pages 519–530.

Haykin, S. (1999). Neural networks a comprehensive introduction.

Heilman, K. M., Nadeau, S. E., and Beversdorf, D. O. (2003). Creative innovation: possible brain mechanisms. *Neurocase*, 9(5):369–379.

Hendriks, P. H. and Fruytier, B. G. (2014). Ordering disorders; linking organization design and knowledge integration. *Knowledge Management Research & Practice*, 12(1):48–61.

Hislop, D. (2013). *Knowledge management in organizations: A critical introduction*. Oxford University Press.

Holland, J. H. (1975). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence.* U Michigan Press.

Holme, P. and Newman, M. E. (2006). Nonequilibrium phase transition in the coevolution of networks and opinions. *Physical Review E*, 74(5):056108.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1992). Geographic localization of knowledge spillovers as evidenced by patent citations. Technical report, National Bureau of Economic Research.

Jones, B. (2009). The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *Review of Economic Studies*, 76(1):283–317.

Kalkan, V. D. (2011). Organizational intelligence: Antecedents and consequences. *Journal of Business & Economics Research (JBER)*, 3(10).

Katila, R. and Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of management journal*, 45(6):1183–1194.

Katzenbach, J. R. (1993). *The wisdom of teams: Creating the high-performance organization*. Harvard Business Press.

Kauffman, S. (1993). *The origins of order: Self-organization and selection in evolution*. Oxford University Press, USA.

Kauffman, S. and Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of theoretical Biology*, 128(1):11–45.

Kim, J., Lee, S. J., Marschke, G., et al. (2004). Research scientist productivity and firm size: evidence from panel data on inventors. *mimeograph, SUNY Albany*.

Kleinbaum, A., Stuart, T., and Tushman, M. (2008). *Communication (and coordination?) in a modern, complex organization*. Harvard Business School Boston, MA.

Kleinberg, J. M. (2000). Navigation in a small world. *Nature*, 406(6798):845–845.

Kogut, B. and Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization science*, 3(3):383–397.

Kogut, B. and Zander, U. (1996). What firms do? coordination, identity, and learning. *Organization science*, 7(5):502–518.

Kossinets, G. and Watts, D. (2006). Empirical analysis of an evolving social network. *Science*, 311(5757):88–90.

Kurzweil, R. (2012). *How to create a mind: The secret of human thought revealed*. Penguin.

Landau, M. (1969). Redundancy, rationality, and the problem of duplication and overlap. *Public Administration Review*, pages 346–358.

Lazer, D. and Friedman, A. (2007). The network structure of exploration and exploitation. *Administrative Science Quarterly*, 52(4):667–694.

Leonard, D. and Straus, S. (1997). Putting your company's whole brain to work. *Harvard Business Review*, 75:110–122.

Levinthal, D. (1997). Adaptation on rugged landscapes. *Management science*, 43(7):934–950.

Levitan, B., Lobo, J., Kauffman, S., and Schuler, R. (1999). Optimal organizational size in a stochastic environment with externalities. Santa Fe Institute.

Lex, R., Kovacs, B., and Vicsek, A. (2011). A comparison of email networks and off-line social networks: A study of a medium-sized bank.

Love, J. H. and Roper, S. (2009). Organizing innovation: complementarities between cross-functional teams. *Technovation*, 29(3):192–203.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization science*, 2(1):71–87.

March, J. G. and Sutton, R. I. (1997). Crossroads-organizational performance as a dependent variable. *Organization science*, 8(6):698–706.

Marengo, L., Dosi, G., Legrenzi, P., and Pasquali, C. (2000). The structure of problem-solving knowledge and the structure of organizations. *Industrial and Corporate Change*, 9(4):757–788.

Martinelli, A. (2011). An emerging paradigm or just another trajectory? understanding the nature of technological changes using engineering heuristics in the telecommunications switching industry. *Research Policy*.

Mason, W. and Watts, D. J. (2012). Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, 109(3):764–769.

Mayr, E. et al. (1963). Animal species and evolution. *Animal species and their evolution.*

Milgram, S. (1967). The small world problem. *Psychology today*, 2(1):60–67.

Mintzberg, H. (1993). *Structure in fives: Designing effective organizations.* Prentice-Hall, Inc.

Moorman, C. and Miner, A. S. (1998). Organizational improvisation and organizational memory. *Academy of management Review*, 23(4):698–723.

Morrison, E. (2002). Newcomers'relationships: The role of social network ties during socialization. *Academy of management Journal*, 45(6):1149–1160.

Mowery, D. C., Oxley, J. E., and Silverman, B. S. (1998). Technological overlap and interfirm cooperation: implications for the resource-based view of the firm. *Research policy*, 27(5):507–523.

Mumford, M. D. and Gustafson, S. B. (1988). Creativity syndrome: Integration, application, and innovation. *Psychological bulletin*, 103(1):27.

Nahapiet, J. and Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. *Academy of management review*, pages 242–266.

Nelson, R. R. and Winter, S. G. (2009). *An evolutionary theory of economic change*. Harvard University Press.

Newman, M. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409.

Newman, M. (2004). Who is the best connected scientist? a study of scientific coauthorship networks. *Complex networks*, pages 337–370.

Newman, M. (2010). *Networks: an introduction*. Oxford University Press, Inc.

Nohria, N. and Gulati, R. (1996). Is slack good or bad for innovation? *Academy of management Journal*, 39(5):1245–1264.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization science*, 5(1):14–37.

Nonaka, I. and Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford university press.

Nonaka, I. and Takeuchi, H. (1997). The knowledge-creating company. *The economic impact of knowledge*, page 183.

Owen-Smith, J. and Powell, W. (2004a). Knowledge networks as channels and conduits: The effects of spillovers in the boston biotechnology community. *Organization Science*, 15(1):5–21.

Owen-Smith, J. and Powell, W. W. (2004b). Knowledge networks as channels and conduits: The effects of spillovers in the boston biotechnology community. *Organization science*, 15(1):5–21.

Palomeras, N. and Melero, E. (2010). Markets for inventors: learning-by-hiring as a driver of mobility. *Management Science*, 56(5):881–895.

Pepe, A. (2011). The relationship between acquaintanceship and coauthorship in scientific collaboration networks. *Journal of the American Society for Information Science and Technology*, 62(11):2121–2132.

Perry-Smith, J. (2006). Social yet creative: The role of social relationships in facilitating individual creativity. *Academy of Management Journal*, 49(1):85–101.

Phelps, C., Heidl, R., and Wadhwa, A. (2012). Knowledge, networks, and knowledge networks a review and research agenda. *Journal of Management*, 38(4):1115–1166.

Podolny, J. (2001). Networks as the pipes and prisms of the market1. *American Journal of Sociology*, 107(1):33–60.

Podolny, J. and Baron, J. (1997). Resources and relationships: Social networks and mobility in the workplace. *American sociological review*, pages 673–693.

Powell, W. W., Koput, K. W., and Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative science quarterly*, pages 116–145.

Quinn, J. B. (1985). Managing innovation: controlled chaos. *Harvard business review*, 63(3):73–84.

Radner, R. (1993). The organization of decentralized information processing. *Econometrica: Journal of the Econometric Society*, pages 1109–1146.

Reagans, R. and McEvily, B. (2003). Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly*, 48(2):240–267.

Reagans, R. and Zuckerman, E. (2001). Networks, diversity, and productivity: The social capital of corporate r&d teams. *Organization science*, 12(4):502–517.

Ridley, M. and Wolf, S. (1998). The origins of virtue: Human instinct and the evolution of cooperation. *Integrative Physiological and Behavioral Science*, 33(1):82–83.

Rivera-Alba, M., Vitaladevuni, S. N., Mishchenko, Y., Lu, Z., Takemura, S.-y., Scheffer, L., Meinertzhagen, I. A., Chklovskii, D. B., and de Polavieja, G. G. (2011). Wiring economy and volume exclusion determine neuronal placement in the drosophila brain. *Current Biology*, 21(23):2000–2005.

Rivkin, J. W. and Siggelkow, N. (2003). Balancing search and stability: Interdependencies among elements of organizational design. *Management Science*, 49(3):290–311.

Sampson, R. (1988). Local friendship ties and community attachment in mass society: A multilevel systemic model. *American Sociological Review*, pages 766–779.

Schumpeter, J. (1939). *Business cycles*, volume 1. Cambridge Univ Press.

Sherif, K. and Xing, B. (2006). Adaptive processes for knowledge creation in complex systems: The case of a global it consulting firm. *Information & Management*, 43(4):530–540.

Simon, H. (1962). The architecture of complexity. *Proceedings of the american philosophical society*, pages 467–482.

Simon, H. A. (1985). What we know about the creative process. *Frontiers in creative and innovative management*, 4:3–22.

Simon, H. A. (2002). Near decomposability and the speed of evolution. *Industrial and corporate change*, 11(3):587–599.

Simon, H. A. and Ando, A. (1961). Aggregation of variables in dynamic systems. *Econometrica: journal of the Econometric Society*, pages 111–138.

Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management science*, 51(5):756–770.

Singh, J. (2008). Distributed r&d, cross-regional knowledge integration and quality of innovative output. *Research Policy*, 37(1):77–96.

Singh, J., Hansen, M. T., and Podolny, J. M. (2010). The World Is Not Small for Everyone: Inequity in Searching for Knowledge in Organizations. *Management Science*, 56(9):1415–1438.

Skellett, B., Cairns, B., Geard, N., Tonkes, B., and Wiles, J. (2005). Maximally rugged nk landscapes contain the highest peaks. In *Proceedings of the 2005 conference on Genetic and evolutionary computation*, pages 579–584. ACM.

Sorenson, O., Rivkin, J., and Fleming, L. (2006). Complexity, networks and knowledge flow. *Research Policy*, 35(7):994–1017.

Spearman, C. (1927). The abilities of man.

Staber, U. and Sydow, J. (2002). Organizational adaptive capacity a structuration perspective. *Journal of Management Inquiry*, 11(4):408–424.

Stephenson, N. (2011). Innovation starvation. *World Policy Journal*, 28(3):11–16.

Stevenson, W. B. (1990). Formal structure and networks of interaction within organizations. *Social Science Research*, 19(2):113–131.

Sytch, M. and Tatarynowicz, A. (2013). Exploring the locus of invention: The dynamics of network communities and firms' invention productivity. *Academy of Management Journal*, pages amj–2011.

Taylor, A. and Greve, H. R. (2006). Superman or the fantastic four? knowledge combination and experience in innovative teams. *Academy of Management Journal*, 49(4):723–740.

Thayer, L. O. and Barnett, G. A. (1996). *Organization–communication: The renaissance in systems thinking*, volume 5. Greenwood Publishing Group.

Tortoriello, M. (2005). The social underpinnings of absorptive capacity: External knowledge, social networks, and individual innovativeness. *Social Networks, and Individual Innovativeness*.

Tortoriello, M. and Krackhardt, D. (2010). Activating cross-boundary knowledge: the role of simmelian ties in the generation of innovations. *Academy of Management Journal*, 53(1):167–181.

Treviño, L., Webster, J., and Stein, E. (2000). Making connections: Complementary influences on communication media choices, attitudes, and use. *Organization Science*, 11(2):163–182.

Tsai, W. (2000). The formation of intraorganizational (linkages. *Strategic Management Journal*, 21(9):925–939.

Tsai, W. (2001). Knowledge transfer in intraorganizational networks: effects of network position and absorptive capacity on business unit innovation and performance. *Academy of management journal*, 44(5):996–1004.

Tsai, W. and Ghoshal, S. (1998). Social capital and value creation: The role of intrafirm networks. *Academy of management Journal*, 41(4):464–476.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American sociological review*, pages 674–698.

Uzzi, B. and Spiro, J. (2005). Collaboration and creativity: The small world problem. *American journal of sociology*, 111(2):447–504.

Von Foerster, H. (1984). *Principles of self-organization—in a socio-managerial context*. Springer.

Walsh, J. P. and Ungson, G. R. (1991). Organizational memory. *Academy of management review*, 16(1):57–91.

Wasserman, S. and Faust, K. (1994). *Social network analysis: Methods and applications*, volume 8. Cambridge university press.

Watts, D. J. (1999). *Small worlds: the dynamics of networks between order and randomness*. Princeton university press.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442.

Weick, K. E. and Roberts, K. H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative science quarterly*, pages 357–381.

Weick, K. E., Sutcliffe, K. M., and Obstfeld, D. (2008). Organizing for high reliability: Processes of collective mindfulness. *Crisis management*, 3:81–123.

Wellman, B., Salaff, J., Dimitrova, D., Garton, L., Gulia, M., and Haythornthwaite, C. (1996). Computer networks as social networks: Collaborative work, telework, and virtual community. *Annual review of sociology*, pages 213–238.

Williams, R. (2006). Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata Journal*, 6(1):58–82.

Williams, W. M. and Sternberg, R. J. (1988). Group intelligence: why some groups are better than others. *Intelligence*, 12(4):351–377.

Woodman, R., Sawyer, J., and Griffin, R. (1993). Toward a theory of organizational creativity. *Academy of management review*, pages 293–321.

Woolley, A. W. and Fuchs, E. (2011). Perspective-collective intelligence in the organization of science. *Organization Science*, 22(5):1359–1367.

Yayavaram, S. and Ahuja, G. (2008). Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability. *Administrative Science Quarterly*, 53(2):333–362.

Yayavaram, S. and Chen, W.-R. (2013). Changes in firm knowledge couplings and firm innovation performance: The moderating role of technological complexity. *Strategic Management Journal*.

Zaidman, N. and Brock, D. M. (2009). Knowledge Transfer Within Multinationals and Their Foreign Subsidiaries: A Culture-Context Approach. *Group & Organization Management*, 34(3):297–329.