# THE TECHNOLOGICAL DISTANCE FROM PARTNERS TO JOINT INVENTION AND ITS EFFECT ON THE VALUE OF ALLIANCE OUTCOME. A RELATIVE ABSORPTIVE CAPACITY PERSPECTIVE

**HUGO ERNESTO MARTÍNEZ ARDILA**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER**
**FACULTAD DE INGENIERÍAS FISICOMECÁNICAS**
**ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE TELECOMUNICACIONES- E3T**
**DOCTORADO EN INGENIERÍA**
**BUCARAMANGA**
**2016**

# THE TECHNOLOGICAL DISTANCE FROM PARTNERS TO JOINT INVENTION AND ITS EFFECT ON THE VALUE OF ALLIANCE OUTCOME. A RELATIVE ABSORPTIVE CAPACITY PERSPECTIVE


**HUGO ERNESTO MARTÍNEZ ARDILA**


**Proyecto de grado para optar al título de Doctor en Ingeniería**


**DIRECTOR:**
**JAIME ALBERTO CAMACHO PICO**
**PhD en Ingeniería Industrial**

**Co-DIRECTOR:**
**ASTRID JAIME ARIAS**
**PhD en Ingeniería Industrial**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER**
**FACULTAD DE INGENIERÍAS FISICOMECÁNICAS**
**ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y DE TELECOMUNICACIONES- E3T**
**DOCTORADO EN INGENIERÍA**
**BUCARAMANGA**



**2016**

*To my family, Mafe and Santi. They have been in the "lows" and "highs", and have thought me the kind of knowledge that goes beyond the boundaries of theories and numbers*

*"For all is like an ocean, all flows and connects; touch it* in one place and it *echoes at the other end of the world." Fyodor Dostoyevsky, The Brothers Kamarazov*

## ACKNOWLEDGEMENTS

I would like to thank the people that had made this doctoral thesis possible and to whom I am greatly indebted. I would like to thank my advisor, Professor Jaime Alberto Camacho Pico, for the support on every aspect of the doctoral process during these hardworking years, and for providing me with the opportunity to show what I can achieve. Likewise, I would like to thank my co-advisor, Professor Astrid Jaime, for her support, backing, guidance, and the many lively discussions about the topics of the thesis in order to make this dissertation a reality.

I would also like to thank the research group Innotec. I was fortunate to be surrounded by a wonderful circle of people from students, professionals, and professors that support me during these years. Thanks to my “dear bro” Gerardo Angulo, doctorate mate with whom I shared lot of thoughts during these years. Thanks as well to the professors from the E3T that helped me in the doctorate endeavours. Thanks to the people that were next to me in the University of Delaware: to the professor Rodrigo Arce; but also to Ana, Juan Camilo, Julian, Edgar, and especially to César. Thanks to Colciencias for the opportunity and the financial support via the scholarship which is gratefully acknowledge.

Finally, thanks to the closest ones. Thanks to my wife for her patience, understanding, but specially her encouragement. My father Hugo and my mother Gloria for their pushing and optimism necessary to finalizing the dissertation. Thanks to my brother Henry, for his help when I needed it; and thanks to my sister Sonia for her positive flow of energy. Thanks to all that people behind the scenario that cannot be perceived although have a fundamental role in this doctorate.

***Hugo***

## CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF APPENDICES

# RESUMEN

**TITULO**: LA DISTANCIA TECNOLÓGICA DE LOS SOCIOS A LA INVENCIÓN CONJUNTA Y SU EFECTO EN EL VALOR DE LOS RESULTADOS DE LA ALIANZA. UNA PERSPECTIVA DESDE LA CAPACIDAD DE ABSORCIÓN RELATIVA*

**AUTOR**: HUGO ERNESTO MARTÍNEZ ARDILA**

**PALABRAS CLAVE**: Capacidad de absorción relativa, valor de innovación, distancia tecnológica, capital tecnológico, alianzas inter-organizativas.

La presente tesis estudia el fenómeno de las alianzas estratégicas y el desempeño de la innovación desde la Visión Basada en el Conocimiento de la firma. Específicamente, la tesis examina como el conocimiento de gran valor puede obtenerse de las alianzas estratégicas. Basándose en la perspectiva de la capacidad de absorción relativa del conocimiento, la tesis analiza el rol de las firmas en las díadas de aprendizaje en las alianzas, como estudiantes o como profesores, y su efecto en el valor de la invención conjuntamente desarrollada. Específicamente, se sugiere que dos atributos importantes, a saber, las distancias tecnológicas relativas y el capital tecnológico de las firmas, son determinantes del valor de la invención conjunta resultante de la alianza. Los métodos utilizados en el estudio incluyen la revisión de la literatura y la investigación empírica con análisis cuantitativo mediante el examen estadístico. La tesis se centra en la industria de la biotecnología y utiliza información de patentes conjuntas extraída de la Oficina de Patentes y Marcas de los Estados Unidos durante los años 2006-2010. Como resultado, la investigación prueba dos factores críticos para obtener innovaciones conjuntas de alto valor. En primer lugar, la importancia de la localización de la base de conocimiento de la invención conjunta en el espacio tecnológico que origina las distancias tecnológicas relativas; y segundo, el rol del capital tecnológico de las firmas en el contexto de la díada de aprendizaje, diferenciando entre las empresas estudiantes y profesores.

---

* Trabajo de grado

** Facultad de Ingenierías Físico Mecánicas Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones. Doctorado en Ingeniería Director: Jaime Alberto Camacho Pico Co-Director: Astrid Jaime Arias

# ABSTRACT

**TITLE**: THE TECHNOLOGICAL DISTANCE FROM PARTNERS TO JOINT INVENTION AND ITS EFFECT ON THE VALUE OF ALLIANCE OUTCOME. A RELATIVE ABSORPTIVE CAPACITY PERSPECTIVE[*]

**AUTHOR**: HUGO ERNESTO MARTÍNEZ ARDILA[**]

**KEYWORDS:** relative absorptive capacity, innovation value, technological distance, technological capital, inter-organizational.

The present thesis deals with the phenomenon of strategic alliances and innovation performance from the Knowledge Based View of the firm. Specifically, the thesis examines how much high valuable knowledge can be obtained from strategic alliances. Based on the perspective of relative absorptive capacity of knowledge, the thesis analyses the role of firms in alliance learning dyads, as students or as teachers, and its effect on the value of their jointly developed invention. Specifically, it is suggested that two important attributes, namely relative technological distances and technological capital of firms, are determinants of the value of the joint invention resulting from the alliance. The methods used in the study include literature review and empirical research with quantitative analysis by means of statistical examination. The thesis focuses on the biotechnology industry and uses joint patents information extracted from the United States Patent and Trademark Office during the years 2006-2010. As a result, the research tests two critical factors to obtain high valuable join innovations. First, the importance of location of the knowledge base of the joint invention in technological space that originates the relative technological distances; and second, the role of technological capital of firms in the learning dyad context differentiating between teacher and student firms.

[*] Work of grade

[**] Faculty of Mechanical Physical Engineering School of Electrical, Electronics and Telecommunications Engineering. Doctorate in Engineering Director: Jaime Alberto Camacho Pico Co-Director: Astrid Jaime Arias

# INTRODUCTION

## Overview

Innovation has been extensively acknowledged as a central element to achieve competitive advantage. In nowadays fast changing environments, firms are increasingly embedded on complex collaborative practices to attain new competences (Hagedoorn, 2002; Laursen & Salter, 2006) making them more dependent on external knowledge in order to develop innovations (von Hippel, 1988). This paradigm has made firms be open (Chesbrough, 2003) and thus dependent, not only on their own capacities, but on external knowledge and experience from other actors (Caloghirou, Kastelli, & Tsakanikas, 2004).

A recognized organizational form used as a mechanism to appropriate and exploit external knowledge to reach and sustain competitive advantages are alliances (Grant, 1996; Kale & Singh, 2000; Simonin, 2004). Besides reasons related to market power, risk sharing, and cost reducing; the alliances theory has been studied by academics and practitioners (Duysters & de Man, 2003; Grant & Baden-Fuller, 2004) thanks to attributes related to complementarities and interdependence between firms (Nohria & Garcia-Pont, 1991; Pfeffer & Nowak, 1976). Nevertheless, although alliances are widely accepted as a mechanism to enhance innovation performance of firms, this matter has received limited support in empirical research (C. Lin, Wu, Chang, Wang, & Lee, 2012).

In alliances, firms are assumed to improve capabilities, resources and competences through transfer, sharing and acquisition of knowledge (Grant & Baden-Fuller, 2004; Lane, Salk, & Lyles, 2001). As a consequence, alliances have a double standpoint that need to be approached: first, their role as a mechanism to acquire knowledge; and second, their role as a mechanism to create knowledge. This phenomenon is found in the research literature in terms of how firms acquire knowledge in alliances to achieve high innovation performance (Bierly, Damanpour, & Santoro, 2009; Camisón & Forés, 2010; de Jong & Freel, 2010; Laursen, Leone, & Torrisi, 2010). A central factor that influences the match between acquisition and creation of knowledge is the firm's ability to acquire, assimilate and exploit external knowledge. This ability is known as the absorptive capacity of firms (W. M. Cohen & Levinthal, 1990) which has been proved to facilitate the inter-organizational collaboration processes (Lane et al., 2001; Mowery, Oxley, & Silverman, 1996).

At the alliance level, this absorptive capacity does not consider only individual attributes but relative characteristics of the firms involved in the collaboration (Lane & Lubatkin, 1998). On the one hand, from this relative absorptive capacity point of view, firms are categorized as teacher and student in the learning processes, a framework that Lane & Lubatkin called "*the learning dyad*". At the same time, the absorptive capacity is path dependent and thus rely on the knowledge or technology cumulativeness of firms (W. M. Cohen & Levinthal, 1990). Hence, the present research argues that the relative characteristics of technological capital of firms, as teachers or students, have an effect on the value of joint innovation in the alliance. This differentiation between the teacher's technological capital and the student's technological capital and their effects on innovation value has been almost completely ignored by the literature since the seminal study of Lane & Lubatkin (1998).

On the other hand, the literature on relative absorptive capacity extensively relies on the technological knowledge similarities or distance between firms in order to study collaboration and knowledge creation (Ahuja & Lampert, 2001; M. J. Benner & Tushman, 2003; Gilsing, Nootebomm, Vanhaverbeke, Duysters, & van der Oord, 2008; Katila & Ahuja, 2002; Nooteboom, Van Haverbeke, Duysters, Gilsing, & Van

Den Oord, 2007). As a consequence, there is an over focus on the study of technological profiles of firms in order to understand how their knowledge bases would affect innovation performance. Nevertheless, the present research argues that the knowledge base reflected on the joint invention, developed by the firms in the alliance, needs to be included as a critical element in the formulae to obtain high value innovations. As a result, a new notion called “relative technological distances” is introduced. These relative distances, which are measured from joint invention to firms’ knowledge bases, are incorporated as an important determinant of innovation value in alliances. What's more, it’s argued that the effect of these relative technological distances also depends on the role of firms in the learning dyad. Consequently, there are two relative distances: the teacher firm technological distance, and the student firm technological distance.

Overall, the present research examines the development and exploitation of knowledge at the outbound limits of the firm (Grant & Baden-Fuller, 2004) in order to achieve the primary goal: the application of existing knowledge to the production of goods and services (Grant, 1996). Accordingly, firms should seek for specialization in knowledge creation and transfer (Kogut & Zander, 1996) because knowledge is the substance that enable the emergence of an efficient and effective innovation (Nonaka & Takeuchi, 1995). The approach where knowledge is the most strategically important resource of the firm is the Knowledge Based View (Grant, 1996, 2013a). Accordingly, the present research is based on the Knowledge Based View of the firm. Besides the mentioned characteristics, this view is appropriate for the nowadays technological changing and turbulent environment in order to obtain competitive advantages (Eisenhardt & Martin, 2000).

**Research Question and Objectives**

Because the importance of innovation, alliances are comprehended as mechanisms through which is possible to achieve and create value. However, this research area needs to be more thoroughly studied and analyzed. The final *purpose* of this research is to improve the competitiveness of strategic alliances by incurring not only on the mechanisms to acquire external knowledge to innovate, but to understand under what conditions is possible to obtain high value knowledge. Therefore, the overarching research question that guides this research is:

*Why are some alliances superior to others in terms of the value of the knowledge created?*

Accordingly, the *main objective* of the research is to develop and test a model about the effect of the technological attributes of the alliance partners on the value of the join invention. The focus is on the characteristics of technological capital, and relative technological distances of firms in the learning dyad framework (Lane & Lubatkin, 1998). This is formulated in the following specific objectives:

In terms of relative technological distances:

i. To determine the effects of the technological distance of the student firm on the value of the joint invention.

ii. To determine the effect of the technological distance of the teacher firm on the value of the joint invention.

iii. To characterize the effect between partner's relative technological distances on the value of the joint invention.

In terms of technological capital:

iv. To determine the effect of the technological capital of the student firm on the value of the joint invention.

v. To determine the effect of the technological capital of the teacher firm on the value of the joint invention.

**Motivation**

This research responds to the statement: “…there are few studies examining the role of R&D alliances in creating new technology based on patent development at the level of individual interaction…” (Lin et al., 2012, pp.283). To help in this aim, the present research is based on the notion of the absorptive capacity of knowledge (W. M. Cohen & Levinthal, 1990). The relative absorptive capacity has revealed that in alliances the firms are dependent on the similarity of knowledge relative attributes of their partners (Lane & Lubatkin, 1998). The absorptive capacity is also dependent of the past cumulative knowledge or, as typically studied, to the past R&D activity. Firms then try to explore new technological domains in a path- dependent way which is a fundamental mechanism to learn and create knowledge (Rosenkopf & Nerkar, 2001). At the alliance level, this technological exploration still operates for each one of the firms. Furthermore, alliances have been studied as mechanisms by which firms explore new technological domains (Rosenkopf & Almeida, 2003). Thus, the present study is aligned to recent research on the argument that the creation of new valuable knowledge in alliances, as mechanisms to acquire knowledge, is reliant on the absorptive capacity of knowledge of the firms involved in it (C. Lin et al., 2012); more specifically, on the relative absorptive capacity between them, a matter barely addressed in the research literature.

An imperative view of the research motivation is related to the global innovation index –GII 2015-(Cornell, INSEAD, & WIPO, 2015). In 2015, Colombia was ranked

67 with a score of 36.41. The Global Innovation Index –GGI– 2015 includes 141 countries that represent the 95.1% of the world's population and 98.6% of the worlds GDP. The GGI includes two sub-indices. First, the Innovation Input Sub-Index with five elements that are critical to enable innovation; and second, the Innovation Output Sub-Index in charge of the results of innovative activities.

The innovation input sub-index has five pillars. In the Business Sophistication pillar, that captures how conducive are the firms to the innovation activity, there are two sub-pillar that are strongly related to this research. They are: Innovation Linkages, and Knowledge Absorption. On the one hand, the Innovation Linkages sub-pillar includes the number of deals on joint ventures and strategic alliances where Colombia is ranked 116. This ranking has been signaled by the study as a weakness indicator in the country. On the other hand, Colombia is ranked 31 in the Knowledge Absorption sub-pillar. This might show that, although the country has medium-high abilities to potentially absorb knowledge, there is a weakness in terms of taking advantage of mechanisms such as strategic alliances. Therefore, improving the understanding about the mechanisms and the processes underlying strategic alliances and absorptive capacity is of vital importance to accelerate the catching up process.

The innovation output sub-index has two pillars: one of them is Knowledge and Technology Outputs, which embraces all those variables that are habitually assumed to be results of inventions and innovations. In the Knowledge and Technology Outputs, one sub-pillar related directly to the present research is included; it is the Knowledge Creation. The knowledge creation sub-pillar includes innovative activities as patent applications and utility models at the national and international levels including even the Patent Cooperation Treaty –PCT–. Colombia is ranked 94 in this sub-pillar, showing again a weakness in terms of knowledge creation in the country. Therefore, nowadays, analyzing how to improve innovation in the country in order to capture and generate wealth is of vital importance.

A central score of the GGI 2015 is the Innovation Efficiency Ratio index. This index is defined as the ratio of the Output Sub-Index to the Input Sub-Index. Colombia is categorized as an inefficient country according to the GII 2015 and is ranked in the 114 place with respect to this ratio. In other words, the resulting outputs are low in the country having into account the current inputs. Therefore, how Colombia can improve its efficiency in terms of innovativeness, is a motivation at present research which brings together two important determinants, knowledge absorption and innovation linkages; and one critical outcome, knowledge creation. These sub-pillars have their equivalents in the present research: alliances context and relative absorptive capacity of knowledge as determinants; and the innovation value as the outcome.

Another reason that motivates the present research is the use of joint patents as indicators of joint innovation activity of firms. Although they are becoming more frequent, keeping a share of 6% of total patents since 1970, firms seem to have an aversion to joint patenting (Hagedoorn, 2003). Thus, it's not easy to find this kind of intellectual property assets (Hagedoorn, van Kranenburg, & Osborn, 2003). Consequently, there is a lack of studies identifying factors that predict the existence of joint patents (C. Kim & Song, 2007). The present research also helps in studying a gap that other studies have declared: "*The phenomenon of joint patenting has been minimally studied and remains–to some degree –an enigma to researchers of innovation and intellectual property*" (Briggs & Wade, 2014; pp.4370). Although, it is an important domain to be researched, most of the literature is centered on examining the quality of joint patents by comparing the joint patents of alliances to the single, monopoly, owned patents (Belderbos, Cassiman, Faems, Leten, & Van Looy, 2014). As a consequence, the present research, in line to some recent studies (Briggs, 2015), focuses on factors that impact the value or quality among the subset of jointly owned patents.

To conclude, the author's personal motivations are also fundamental for the present research. Learning, personal development, the creation of a structured form of thought, and a better understanding of the research methods to solve practical problems are important elements that have been considered (Easterby-Smith, Thorpe, & Jackson, 2008). A personal motivator for this research is the work made by the research team INNOTEC at the Universidad Industrial de Santander in matters related to management of technology (MOT) and management of innovation (MOI); the author has benefited from the team's collaborators', and specially supervisors' experience. Additionally, having participated as a technology coordinator at the technology transfer office *OTRI Estratégica de Oriente*, has motivated the author's research in having a practical view about the technology transfer processes and specially the difficulties and challenges that emerge in terms of the creation of knowledge in strategic alliances among actors in the innovation system.

# 1. THEORETICAL PERSPECTIVES

This study is based on the knowledge based view (KBV) of the firm and the relative absorptive capacity of knowledge (ACAP) perspective. The knowledge based view suggests that firms should be analyzed based on their knowledge resources (Grant, 1996). The absorptive capacity of knowledge is a learning firm's ability to identify, assimilate and exploit external knowledge (W. M. Cohen & Levinthal, 1990) and is directly associated to the firm's prior related knowledge or knowledge base. In the following, these theoretical perspectives are described in more detail focusing to a large degree in their role as key elements of the innovation arena as creators of new knowledge in inter-organizational collaborations or alliances.

## 1.1 THE KNOWLEDGE BASED VIEW (KBV) OF THE FIRM

The Knowledge Based View understands knowledge as the most strategic resource of the firm (Grant, 1996). The present study considers the KBV as a suitable theory due to its ability to clarify the existence of firms as a consequence of their ability to create and transfer knowledge (Kogut & Zander, 1992; Nonaka & Takeuchi, 1995), and make an effective use of it (Rebolledo & Nollet, 2011) in order to obtain competitive advantages (Kogut & Zander, 1992).

The creation and application of knowledge are fundamental processes to the creation of value (J. C. Spender, 1992). Creation and application of new knowledge implies diversity and combination of different types of existing knowledge (Grant & Baden-Fuller, 2004). Therefore, the creation of value is a result of diversity and combination of existing knowledge in organizations. Consequently, organizational forms with a higher potential of combinative knowledge such as inter-firm collaborations are specially analyzed under the lens of the knowledge based view. Furthermore, the knowledge based theory of the firm is better suited to identify settings in which collaboration between firms is higher to market or hierarchical governance in efficiently exploiting and incorporating knowledge (Grant & Baden-Fuller, 1995).

Accordingly, knowledge grounded approaches offer a richer theoretical foundation for the examination of organizational forms such as strategic alliances (Grant, 2013). The term 'strategic alliance' generally denotes agreements of two or more firms to reach a common goal involving their resources and activities (Teece, 1992). However, this thesis focus on the underlying knowledge aspects of the firm and consequently uses the definition of strategic alliances as "*voluntary arrangements between firms with the objective of jointly creating, transferring or applying knowledge to commercial ends*" (Meier, 2011, pp. 3 ). Thus, in this study alliances operate as conduits for inter-organizational creation, transfer and application of knowledge (Kale & Singh, 2000; Lane & Lubatkin, 1998; Lane et al., 2001; Mowery et al., 1996; Simonin, 2004; Stellner, 2015; Zidorn & Wagner, 2013).

Strategic alliances involve a range of collaborative practices that include supplier-buyer partnerships, outsourcing agreements, technical collaboration, joint research projects, shared new product development, shared manufacturing arrangements, common distribution arrangements, cross-selling arrangements, and franchising (Grant & Baden-Fuller, 2004). It is generally accepted from past studies that the concentration of alliances in R&D intensive sectors points to technology as essential in alliance development (Dickson & Weaver, 1997; Doz, 1988; Hagedoorn, 1993).

The present thesis focus on the type of strategic alliances related to joint research projects where the distinctive outcome is the new technological knowledge.

The perception of knowledge creation is associated to the development of new or novel knowledge (Argote, McEvily, & Reagans, 2003; Teece, 1998). New knowledge is the result of combination of existing knowledge (Grant & Baden-Fuller, 2004). In the alliance context, knowledge creation describes the cooperative development of new knowledge by alliance partners (M. Lubatkin, Florin, & Lane, 2001; Reid, Bussiere, & Greenaway, 2001). Grant (1996) identified that characteristics of the donor firm and the recipient firm in alliances are central to the creation of competitive advantages. Differences in the characteristics of the source of knowledge, characteristics of the recipient of knowledge, and characteristics of the context are key to understand difficulties in the knowledge transfer process (Szulanski, 1996). Furthermore, inter-organizational knowledge outcomes are basically determined by the characteristics of knowledge itself, the alliance partners, and those of their interaction and relationship (Argote et al., 2003; Nooteboom et al., 2007; Simonin, 2004). If individual knowledge characteristics of donor and recipient firms are crucial in the alliance; then, it is apparent that knowledge (e.g. technological, market, management) is relative to each firm in inter-firm collaborations (Sammarra & Biggiero, 2008). Consequently, new jointly created knowledge in strategic alliances might have a relative relation to each of the partners' existing knowledge. This is analogous to thinking that central elements in the knowledge context are related to the properties of knowledge, properties of units, and the relationship of units (Argote et al., 2003). This relative approach is fundamental in the present thesis which argues that the success of alliances is not based only on how much knowledge has been transferred, but how successfully alliance partners have been jointly applying existing knowledge (Meier, 2011).

The knowledge based view suggests that a firm can be described by the knowledge it integrates (Grant & Baden-Fuller, 1995). Knowledge outcomes appear themselves in the firm's knowledge assets (Meier, 2011). Then, knowledge assets such as patents, new products or technologies are a reflection of knowledge creation in alliances (Reid et al., 2001). Thus, a firm in an alliance and the alliance itself may be described by the knowledge assets it possesses such as patents, new products or technologies. The present research thesis describe the firms in an alliance by their patents based on the assumption that patents are outcomes or knowledge assets which reflect the knowledge firms in alliances possesses or integrates.

The inventions embedded in patents have been used in the literature as a proxy to innovations (C. Kim & Song, 2007; C. Lin et al., 2012; Nooteboom et al., 2007). In the knowledge management literature, knowledge is the essence of the innovation process (Nonaka & Takeuchi, 1995; Rodan & Galunic, 2004). Several models of the innovation process usually consider the outcome of the innovation process to be new knowledge, which is subtly paralleled with innovation (Galunic & Rodan, 1998; C. Kim & Song, 2007; C. Lin et al., 2012; W. Tsai & Ghoshal, 1998). Following this reasoning, the present research uses the technological knowledge embedded in joint inventions (e.g. patents) as a reflection of the innovation outcome from the alliance.

Finally, besides the above mentioned characteristics, the absorptive capacity of knowledge is a key factor that impacts positively the knowledge transfer in alliances (Easterby-Smith, Lyles, & Tsang, 2008; van Wijk, Jansen, & Lyles, 2008). The knowledge based view is relevant to the absorptive capacity construct because the absorptive capacity is a crucial element to develop and grow the firm’s knowledge base (Volberda, Foss, & Lyles, 2010). The absorptive capacity describes the organization’s ability to learn from the partner (Steensma & Lyles, 2000). The prevailing literature on absorptive capacity generally confirms the importance of overlap in knowledge in elucidating the joint creation in alliances (Meier, 2011). This element of knowledge overlapping is essential to the argumentation of joint new knowledge creation in this research thesis. The following section will focus on the concept of the absorptive capacity of knowledge as a theoretical base on behalf of the present research study.

## 1.2 THE RELATIVE ABSORPTIVE CAPACITY PERSPECTIVE

The absorptive capacity of knowledge is one of the most studied concepts in technology and innovation management in the last two decades (Martinez, Jaime, & Camacho, 2012). Initially, the term absorptive capacity was defined as the ability to learn from external sources of knowledge (W. M. Cohen & Levinthal, 1989, 1990, 1994). Thus, knowledge is crucial to it. This '*absorptive capacity*' that W. M. Cohen & Levinthal make equivalent to firm '*learning*' (1990) distinguished the twofold part of R&D: firms not only would have the ability to reproduce innovation process and products but also have the ability to exploit external knowledge to create new knowledge (e.g. innovations). Thus, absorptive capacity can be thought as a determinant to innovate. This capacity is a function highly dependent on prior related knowledge or knowledge bases of firms.

There have been numerous definitions and reconceptualization of the absorptive capacity notion after the seminal article of W. M. Cohen & Levinthal (1990). For example, Mowery and Oxley (1995) define it as an ability to manage the tacit nature of knowledge and its transformation; Kim (1993) defines it as the capacity to learn and solve problems to assimilate and create new knowledge; Zahra and George (2002) made a reconceptualization and define it as a "*set of organizational routines and processes by which firms acquire, assimilate, transform, and exploit knowledge*" (pp. 186); and Lane, Koka, & Pathak (2006) propose a new more detailed definition of this ability through three sequential processes: *"(1) recognizing and understanding potentially valuable new knowledge outside the firm through exploratory learning, (2) assimilating valuable new knowledge through transformative learning, and (3) using the assimilated knowledge to create new knowledge and commercial outputs trough exploitative learning*" (pp. 856).

The contributions to the concept of absorptive capacity have had the challenge to highlight it through other ways besides the R&D costs measured. In other words, based on the initial argument of Cohen and Levinthal, firms should incur on substantial costs on R&D in the long term in order to develop a knowledge base which in turn is the absorptive capacity of the firm (1989). Thus, the generalized perception is the idea that firms just need to have strong investment in R&D, assuming that a firm possesses in advance the whole resources and mechanisms, in order to exploit the external knowledge. Lane and Lubatkin (1998) criticize this argument based on the premise that firms with the same level of absorptive capacity (same R&D investment) have not equal capacity to learn from all other organizations. These authors interpreted the concept as a learning dyad construct - a teacher firm and a student firm- and named it Relative Absorptive Capacity. The Relative Absorptive Capacity is the firm’s ability to learn from another firm depending on their similarity of knowledge and organizational practices (Lane & Lubatkin, 1998).

The existing quantitative empirical literature on absorptive capacity largely confirms the importance of knowledge similarity for joint creation and knowledge transfer between firms (Meier, 2011). The present research thesis emphasizes the dimension of similarity of knowledge bases between teacher and student firms in the learning dyad model of the Relative Absorptive Capacity of Knowledge. This remains the discussion about the absorptive capacity in inter-organizational settings where knowledge is possessed by firms and can be abstracted as a tangible asset interacting between organizations (Marabelli & Newell, 2014).

The role of knowledge in the learning process was previously perceived by the seminal work of Cohen & Levinthal (1990) that stated “…some portion of that prior knowledge should be very closely related to the new knowledge to facilitate assimilation, and some fraction of that knowledge must be fairly diverse, although still related, to permit effective, creative utilization of the new knowledge" (pp. 136). In other words, as Kim and Inkped (2005) said "there is a tension between the need for diverse technologies in order for firms to have something to learn from each other and the need for similar technologies to allow the firms to learn from each other" (pp.319). This argument is in line to some studies which are centered in the multidimensional character of knowledge (Lichtenthaler, 2009; Vasudeva & Anand, 2011; Zahra & George, 2002). However, it is in the perspective of relative absorptive capacity where first an interaction analysis between teacher and student firms emphasizes the inter-organizational level (Lane & Lubatkin, 1998) in contrast to the isolated firm analysis of the seminal work of Cohen and Levinthal (1990). Then, it is understandable that the closeness or farness between firms' technological knowledge bases in an inter-organizational context is a significant element in order to learn and create new knowledge.

Firms develop capacities to diversify their knowledge and improve innovation performance (K. H. Tsai, 2009). This argument is especially functional at the inter-organizational level because it allows firms to have a broad access to technological opportunities (Oerlemans & Knoben, 2010). In addition, technological opportunities are localized in technological space. Then, knowledge from firms has a spatial and technological connotation and its absorption can depend of how far or distant it is from a source knowledge base domain (Rosenkopf & Nerkar, 2001). This idea was hinted in Cohen and Levinthal's work in their argument about the addition of efforts by firms when the knowledge domain was not closed; and has been thereof reflected on some studies that from the knowledge similarity or distance dimension of the relative absorptive capacity perspective, have considered the phenomenon of cognitive proximity and technological knowledge distance between firms (Lane et al., 2001; Makri, Hitt, & Lane, 2010; Nooteboom et al., 2007; Quintana-García & Benavides-Velazco, 2010; Sapienza, Parhankangas, & Autio, 2004; Shin & Jalajas, 2010).

The present thesis makes use of the knowledge technological distance notion in order to localize firms' knowledge and alliances new jointly created knowledge in a technological space. These concepts help to analyze and improve the understanding about how alliance knowledge characteristics in a technological space affect their own innovation performance outcomes. In the next sections, a review on the technological distance concept is made focusing on the ambiguity of different definitions, a definition proposal, its measurement, and its role in the innovation domain on the inter-organizational context.

# 2. THE CONCEPT OF TECHNOLOGICAL DISTANCE

In general, it is understood that technological distance examines the differences between the technological knowledge and expertise used by firms (Enkel & Gassmann, 2010; Gilsing et al., 2008; C. Lin et al., 2012; Schulze & Brojerdi, 2012). However, technological distance also reflects the notion that ideas are more or less related (McNamee, 2013). Because of that, some studies understand technological distance by the degree of technology or knowledge '*overlap*' between the firms in alliances–the more the overlap, the less the distance (Cowan & Jonard, 2008; Laursen et al., 2010; Nambisan, 2013). In this respect, in the literature is usual to find the use of other terms or jargon to explain the same idea of technological distance (e.g. technological proximity, technological similarity, and technological relatedness). In the following paragraphs, the relationships between the terms are explained to improve the understanding of why they are used in an interchangeable way as the technological distance concept.

## 2.1 KEY INTERRELATED CONCEPTS

**2.1.1 Technological proximity** Technological distance and technological proximity are terms with opposite measures of the same concept. The longer the technological distance between two firms, the less their technological proximity and vice versa. The two concepts appear in the literature as the two sides of the same coin, and because of that, they have been used interchangeably but with some differences in the interpretation of their results (Rosenkopf & Almeida, 2003). In general, technological distance refers to the absence of overlap between knowledge bases of the firms involved (Cowan & Jonard, 2008; Laursen et al., 2010; van de Vrande, Vanhaverbeke, & Duysters, 2011); and technological proximity refers to the knowledge base overlap between the firms (J. Cantwell & Colombo, 2000; Mowery, Oxley, & Silverman, 1998; Schoenmakers & Duysters, 2006). Consequently, both of them are strongly concerned with the idea of overlapping of knowledge bases among actors; i.e. more overlap between firms, more proximity and less distance between them (Stellner, 2014). In our analyses, therefore, higher values of technological proximity connote lower technological distance, and vice versa.

**2.1.2 Technological similarity** Accordingly, '*knowledge overlap*' is seen as a negative measure of knowledge distance -and a positive measure of knowledge proximity- and is understood as the number of facts which both firms know or have in common (Cowan & Jonard, 2008). Knowledge overlap can be defined as the degree to which firms possess '*similar*' types of knowledge and information (Schulze & Brojerdi, 2012). Therefore, by examining their technological overlap, it is possible to understand how similar or distant in technology the firms are (Schildt, Keil, & Maula, 2012). Furthermore, knowledge proximity represents the similarity of knowledge bases and shared understanding (Dangelico, Garavelli, & Petruzzelli, 2010; Mattes, 2012). In view of that, some studies use knowledge similarity and knowledge proximity as the same idea (Boschma, 2005), both of them are associated to the degree of overlap between actors' knowledge bases (Fornahl, Broekel, & Boschma, 2011). In the present analysis higher values of technological distance connote lower similarity, and vice versa (Rosenkopf & Almeida, 2003).

**2.1.3 Technological relatedness** *'Overlap'* also indicates the extent to which firms have related knowledge antecedents (C. Kim & Song, 2007; Mowery et al., 1996). Technological or knowledge relatedness describes the extent of similarity and compatibility of technology or knowledge between two individuals or organizations (Lane & Lubatkin, 1998). Consequently, '*knowledge relatedness*' has been defined also as the degree of overlapping among knowledge bases between firms (Petruzzelli, 2011; Sapienza et al., 2004) which is in line with the mentioned concepts of '*similarity*' and '*proximity*'. Nevertheless, some studies offer a wider concept of knowledge relatedness, from which knowledge proximity is only a category of dimensions among other categories such as commonality -same type of knowledge used in more than one technology- and complementarity -need to use together different technologies- (Breschi, Lissoni, & Malerba, 2003). However, relatedness has commonly been defined in broad terms, often using similarity and complementarity interchangeably (Davis, Robinson Jr, Pearce II, & Park, 1992; Farjoun, 1998).

**2.1.4 Technological distance and cognitive distance** Similarly, the concepts found in the literature of '*cognitive distance'* and '*technology distance*' -or cognitive proximity and technology proximity- are used without distinction. For example, Broekel & Boschma (2012) refer to cognitive proximity as the technological similarity of two organizations' knowledge bases. This conceptual intersection can be better understood if the general definition of cognitive proximity -or distance- is taken into consideration: cognitive proximity is concerned to the similarities (Boschma, 2005; Dangelico et al., 2010) or the degree of overlap between the actors' knowledge bases (Broekel & Boschma, 2012; Buerger & Cantner, 2011; Fornahl et al., 2011). Cognitive distance is then concerned with the differences in the actors' knowledge bases and expertise (Enkel & Gassmann, 2010; Gilsing et al., 2008; Hautala, 2011). Consequently, the definition of cognitive distance and technological distance seem to be highly related and are based on the same principles of knowledge overlap and knowledge similarity already explained in this section.

Formally, the notion of cognitive distance is based on an interactionist view of knowledge, to capture differences in knowledge and skills or cognitive frames between two entities (Weick, 1995). This differences are captured in the way actors perceive, interpret, understand and evaluate the world (Wuyts, Colombo, Dutta, & Nooteboom, 2005). Nooteboom (2000) defines cognitive distance in terms of common knowledge base and expertise among actors. Cognitive distance -or proximity- is close to the concept of technology distance –or proximity-; however the latter is based on the similarity among actors in terms of only technology knowledge bases (Schamp, Rentmaister, & Lo, 2004). From this point of view, the technological distance is seen as a part of one of the multiple dimensions (e.g. marketing, organizational, technological) that are embraced by the cognitive distance concept (Cunningham & Werker, 2012; Hautala, 2011; Nambisan, 2013).

Other studies find another fundamental difference between these two notions. They argue that cognitive proximity is broader and refer to '*how*' actors interact, whereas technological proximity refers to '*what*' they exchange and the potential value of these exchanges (Knoben & Oerlemans, 2006a; Petruzzelli, Albino, & Carbonara, 2009). However, if technology are the tools, devices and knowledge that mediate between inputs and outputs and that create new products or services (Tushman & Anderson, 1986); technological proximity refers not only to these technologies themselves, but from the Knowledge Based View, to the knowledge actors possess about these technologies (Knoben & Oerlemans, 2006a; Mowery et al., 1998; Petruzzelli, 2011). In addition, in some research works the cognitive distance is specified in terms of technological distance for various reasons including that the studies normally use patents in order to measure innovative success where the technological knowledge is dominant, and because it is not clear how to measure other dimensions of cognitive distance (Gilsing et al., 2008; Wuyts et al., 2005). In conclusion, the technological distance is also used as a proxy variable of cognitive distance as an approach to understanding how technological knowledge bases overlapping occurs among actors.

**2.1.5 Technological distance and technological space** The literature also makes use of the concept of technological space when talking about technological distance. In these studies, firms or entities are localized in a multidimensional technological or knowledge space (M. Benner & Waldfogel, 2008; Jaffe, 1986;

Olsson & Frey, 2002) by positioning their technological knowledge parts (e.g scientist, patents, clusters) and setting them in relation to other companies or entities technological knowledge parts (Stellner, 2014).

Technological space might be considered as the realm made up of all human knowledge or inventive activity (M. Benner & Waldfogel, 2008). This space is supposed to be fully connected and continuous. Therefore, two entities, technologies or firms in technological space can share some extent of similarity or have a measurable distance between them (McNamee, 2013). The distance measure offers the idea that each firm has a location in an N-dimensional knowledge space, where the N-dimension might represent a unique technological class or area (A. G. Z. Hu & Jaffe, 2003). Because it is impossible to observe directly the knowledge of a firm, the use of patent data offers a convenient way into its knowledge; this might be a good reason why patents are used in most of the studies (M. Benner & Waldfogel, 2008).

Some studies go further and define the technology space as a metric space where the body of technology is regarded as a set. In this space ideas are separated by technological distance, and the ideas contained in the technology set form an infinite, bounded, closed and connected set (Olsson & Frey, 2002). In this model, the technological distance is a real-valued metric that defines ideas relative position

in technology space. For instance, to obtain the distance is important to define the dimensions of the space; and firms need to be positioned in technological space. Typical distance measures assume that dimensions in technological space are unrelated; then “overlap” in the same dimension of knowledge space have an increase in proximity as was stated before (Stellner, 2014).

## 2.2 DEFINING TECHNOLOGICAL DISTANCE: A PROPOSAL

The Oxford dictionary[1] defines distance as "*the length of the space between two points*" or "*the avoidance of familiarity*"[2]. Two comparisons can be made having into account this definition to delineate an inclusive concept of technological distance. First, based on the explanations made in the above sections, it is understandable that the "two points" included in the Oxford definition can be related to the two firms, actors or technologies under observation; and that "the length of the space" indicates how far apart or how close these firms are. Second, if the goal is to elucidate the concept of technological distance, then the two firms need to be positioned in a technological space; and therefore, their technologies might give an indication of how far apart/close to, or how familiar/unfamiliar are the firms in this space.

As mentioned, technology is the tools, devices and knowledge that mediate in processes and/or the creation of new products and services (Tushman & Anderson, 1986). However, as stated the approach of the analysis is not centered on the technologies by themselves, but in the knowledge the firms possess about these technologies (Jaffe, 1986; Knoben & Oerlemans, 2006a; Mowery et al., 1998; Petruzzelli, 2011). Therefore, to understand what is the technological distance between two firms, actors, or in general, entities in technological space, is essential to understand how far apart/close to or how unfamiliar/familiar are their respective technological knowledge. In conclusion, the definition of technological

[1] OXFORD DICTIONARIES Distance [on line] [Accessed the 06/JUN/2014] .http://www.oxforddictionaries.com/es/definicion/ingles/distance?q=DISTANCE.

[2] In order to have another meaning of the concept, the definition from the Spanish Royal Academy (RAE for its initials in Spanish) was consulted. The RAE defines distance as "*The space or interval of place or time that mediates between two points or events*" or "*The noteworthy difference or dissimilarity between one thing and another*". These two definitions of the RAE are equivalent to the definitions of the Oxford dictionary.

distance can be composed in a comprehensive way according to the research objective of this study:

*Technological distance is the extent of how far apart/close to or unfamiliar/familiar is the technological knowledge between two entities under observation which are positioned in a technological space.*

This definition might clarify the ambiguity generated by the different terms used to denote the same concept of technological distance (e.g. technological proximity, technological similarity, and technological relatedness). In other words, the use of the idea expressed by the words of '*how far or close to*', which at the same time is based on the idea of '*length*', make that separate used terms in the literature such as distance and proximity converge into the same concept. Hence, the notion of how overlapped are the knowledge bases between the firms, or entities is also aligned to this definition in order to understand how similar, related, proximate or distanced they are.

This definition have some commonalities with the study of Olsson and Frey (2002) who stated that "*technological distance reflects the notion that ideas are more or less related. For instance, the two ideas 'steel' and the 'the Bessemer process' are more closely related than the ideas "the Bessemer process' and the 'spinning wheel'"* (pp.71). Then, technological space in the definition is seen as the whole technological human knowledge, experiences or inventive activity made of ideas which are separated in different grades or lengths (Olsson & Frey, 2002). This space can also be represented in multiple dimensions, where each dimension might characterize classes or fields of technologies (M. Benner & Waldfogel, 2008). This facilitates the location of firms and their knowledge bases in space to operationalize the concept of technological distance; and helps to use tools and listings already set up in different endeavors by academics and practitioners (e.g. patents, human resources specialized in technological fields, etc).

# 3. MEASURING TECHNOLOGICAL DISTANCE

In the literature, technological distance is measured in various ways. In general, technological distance has been measured using two indirect indicators of patent data: vectorized patent data and citation patent data. However, more direct indicators were also developed, although much less used, which are based on interviews and surveys carried out with the organizations. Each of these approaches has advantages and disadvantages.

The main shortcomings of using patent data are related to the arguments that not all knowledge flows are captured by citations (A. G. Z. Hu & Jaffe, 2003) and not all technologies and innovations are patented by firms, for example because of firms' strategic reasons (C. Lin et al., 2012). Also, patents focus only in the codified and explicit dimensions of knowledge, neglecting its tacit nature (Dangelico et al., 2010; Petruzzelli et al., 2009). Different studies are well aware of these drawbacks which have been intensively discussed in the literature (Buerger & Cantner, 2011).

Notwithstanding, patents are the most used data in measuring technological distance and in general innovation. This can be mainly due to their availability and relatively easy access (e.g USPTO, EPO, JPO data bases). Although patents do not explicitly state the technology's '*know-how*', this limitation is attenuated by the argument that codified knowledge flows, reflected in patents, and tacit knowledge flows are linked and complementary (Mowery et al., 1996). Patents provide a set of unique information to analyze inter-organizational alliances characteristics and their innovation process dynamics (Petruzzelli, 2011). These characteristics such as quantity, geography, and technological aspects (Buerger & Cantner, 2011) are key elements embedded in coding conventions which facilitate the study using large samples (Makri et al., 2010) which consequently increase the external validity for relevance in the research.

On the other hand, some more direct indicators were developed to deal with the shortcomings of measuring indirectly innovation and technology using patents. These indicators are essentially related to interviews and survey data. Interviews and surveys are useful methods of collecting direct information about technologies and innovation, and have been used at different levels such as organizations (Bierly et al., 2009; Hautala, 2011), regional (UIS, UNAB, & SETIC, 2013), national (DANE, 2015), and international endeavors (OECD, 2009). These methods use perceptual measures than more objective proxies (e.g. patents) which at the same time tend to be less precise (Grant, 1996). Consequently, these methods have less objectivity than using patents, which results in less external validity. Furthermore, the significance and the representativeness of the results depend widely on the answer rates (Archibugui & Sirilli, 2001). This is in line to the knowledge based view paradox that states the difficulty to measure the concepts of theoretical interest in research (Sapienza et al., 2004; J. Spender & Grant, 1996).

In the following section, the measures are described focusing in four used methods: vectorized patent data, citation patent data, other recent forms of measurement, and interview and survey data. The vectorized patent data comprise the angular separation measure introduce by Jaffe (1986), the Pearson correlation coefficient (M. Benner & Waldfogel, 2008), the correlation of revealed technological advantage –CRTA- (Nooteboom et al., 2007), and the Euclidean distance used by Rosenkoof and Almeida (2003). The citation patent data section describes the citation overlap measure (Stuart & Podolny, 1996), and the cross citation ratio (C. Kim & Song, 2007). Other forms of distances comprise more recent measures such as the mutual information measure (Cunningham & Werker, 2012), the Min- complement distance (Bar & Leiponen, 2012), the Mahalanobis distance (Bloom, Shankerman, & Van Reenen, 2013), the weighted angular separation and the aggregated patent-to-patent angular separation (Stellner, 2014). Finally the interviews and survey data are described in the last sub-section.

### 3.1 VECTORIZED PATENT DATA

- **Angular separation:** Most of the studies use this approach. The analyses are based on the work proposed by Jaffe (1986) who, grounded on the recommendation of Griliches (1979), used the patent technological class information to construct a vector based proximity measure of the closeness between two actors in the technology space (e.g. see Benner & Waldfogel, 2008; Buerger & Cantner, 2011; Greunz, 2003; Hu & Jaffe, 2003; J. W. Kim & Lee, 2004; Lin et al., 2012; Ornaghi, 2009; Petruzzelli, 2011; Reitzig & Wagner, 2010; Scherngell & Barber, 2011, 2009; van de Vrande et al., 2011; Verdolini & Galeotti, 2011). The Jaffe's measure is calculated as follows:

$$Technological\ Similarity_{ij} = \left( \frac{f_i f_j'}{\sqrt{(f_i f_i')(f_j f_j')}} \right) = \left( \frac{\sum_{k=1}^{n} f_{i,k} \cdot f_{j,k}}{\sqrt{\sum_{k=1}^{n} f_{i,k}^2} \sqrt{\sum_{k=1}^{n} f_{j,k}^2}} \right)$$

**Equation 1 Jaffe measure of proximity based on cosine similarity**

Where $fi$ and $fj$ are vectors formed by all the patents registered by firm $i$ and firm $j$ during a period of time and allocated to a specific $k$ patent class. This measure is normally known as the cosine similarity or un-centered Pearson correlation metric which ranges from -1 to 1. In our case, the cosine similarity is used only in the positive technological space (the vectors of firms $fi$ and $fj$ are always positive) making the range limited from 0 to 1. The distance, named cosine distance, results by subtracting the un-centered correlation from 1.

Some studies use variations or extensions from the work of Jaffe (1986). For example, McNamee (2013) uses the cosine similarity based on taxonomical methods with a focus on the hierarchy of a patent classification (e.g. International Patent Classification-IPC-); and Bloom et al. (2013) use a measure similar to the angular separation but without adjustment to the length of vectors.

- **Pearson correlation coefficient**: The Pearson correlation coefficient of the technology vectors of firms $i$ and $j$ (M. Benner & Waldfogel, 2008) is defined as:

$$r_{ij} = \frac{Cov(f_i, f_j)}{SD(f_i) * SD(f_j)}$$

**Equation 2 Pearson correlation coefficient measure**

Where $Cov$ is the covariance between the two vectors, and $SD$ is the standard deviation. Its range is from -1 (large distance) and 1 (high proximity) (Stellner, 2014). A distance metric for the two vectors is known as Pearson's distance which is defined as $dij$=1−$rij$; and its range is from 0 to 2.

Some studies calculate the technological distance based on the squared of the Pearson correlation -coefficient of determination- (Scherngell & Barber, 2009, 2011; Scherngell & Hu, 2011). In these studies dij=1−$r$2, where dij is the technological distance between actors $i$ and $j$, and $r$2 is the squared of the Pearson correlation coefficient between the technological vectors. The range of the coefficient of determination is from 0 to 1.

- **Correlation of revealed technological advantage (CRTA):** the CRTA is the Pearson correlation index of the distribution across technological classes of the Revealed Technological Advantage (RTA) of each firm relative to the other sample firms (Gilsing et al., 2008; Nooteboom et al., 2007; Wuyts et al., 2005). The RTA of a firm in a particular technological field is given by the firm's share in that field or technological class of the patents granted to all companies in the study, relative to its overall share of all patents granted to those companies. Stellner (2014) defines RTA as:

$$RTA_{iu} = \frac{g_{iu} / \sum_{n=1}^{N} g_{nu}}{\sum_{t=1}^{T} g_{it} / \sum_{n=1}^{N} \sum_{t=1}^{T} g_{nt}}$$

**Equation 3 Revealed technological advantage measure**

Where $giu$ is the number of patents owned by firm $i$ in technology class $u$. $N$ is the total of firms, and $T$ is the total of technology classes. The outcome from the RTA calculation is a vector

$$RTA_i = (RTA_{i1}, \dots, RTA_{iT})$$

**Equation 4 Vector of revealed technological advantages**

The correlation between two RTA vectors from two firms $i$,$j$ is then:

$$r_{ij} = \frac{Cov(RTA_i, RTA_j)}{SD(RTA_i) * SD(RTA_j)}$$

**Equation 5 Correlation between the RTA of firms**

For a more detailed information of CRTA see J.A. Cantwell & Barrera (1998) and J. Cantwell & Colombo (2000).

- **Euclidean distance**: Other studies calculate the Euclidean distance between the patent class vectors for each pair of firms (Rosenkopf & Almeida, 2003). In this case the technological distance between firms is defined as:

$$d_{ij} = \sqrt{\sum_{c=1}^{N} \left(p_i^c - p_j^c\right)^2}$$

**Equation 6 Euclidean distance measure**

Where N is the number of dimensions (patent classes) in technological space; and $pic$,$pjc$ are the number of patents in class $c$, or the share from the total patents in that class of firms $i$,$j$ respectively (Ahuja, 2000; Rosenkopf & Almeida, 2003). The Euclidean distance is sensible to the distribution of patents in sparse technological vectors because it includes the non-zero dimensions or technological classes where the firms have not activity.

## 3.2 CITATION PATENT DATA

- **Citation overlap:** this measure is based on the work of Stuart & Podolny (1996) and Mowery et al. (1996). This measure asks how many of the patents that one firm cites are also cited by its partner firm. If the universe of cited patents runs from 1 to $T$, and $\delta_i^j$ is an index that is 1 if firm $j$ cites patent $i$ and 0 otherwise, then the overlap between firm $j$ and $k$ is:

$$\frac{\sum_{i=1}^{T} \delta_i^j \delta_i^k}{\sum_{i=1}^{T} \delta_i^j}$$

**Equation 7 Common cites between firms' measure**

The ratio is the share of the patents cited by firm $j$ that are also cited by firm $k$. This measure is not symmetric because the denominator is the number of patents cited by either firm $k$ or firm $j$ (M. Benner & Waldfogel, 2008), and consequently the similarity is measured from the relative position of one firm when compared to other.

- **Cross-citation ratio:** This cross-citation ratio combine the similarity indicator of both firms by adding them in just one expression:

$$Sim = \left(\frac{firm\ j's\ patent\ cited\ by\ firm\ i's\ patent}{total\ firm\ i's\ citation}\right) + \left(\frac{firm\ i's\ patent\ cited\ by\ firm\ j's\ patent}{total\ firm\ j's\ citation}\right)$$

**Equation 8 Cross citation ratio measure**

The cross-citation ratio captures the extent to which firms $j$ and $k$ cite each other's patents and reflects the degree to which firms $j$ and $k$ are researching in similar technological areas (C. Kim & Song, 2007; Soh, Mahmood, & Mitchell, 2004). The higher the cross-citation ratio, the higher the technical proximity of firms.

### 3.3 OTHER RECENT MEASURES

- **Mutual information:** The mutual information of two random variables is a measure of the variables mutual dependence. It relates to the mathematical theory of communication (Shannon, 1948) and is a fundamental piece in the information theory research area. The mutual information is calculated as (Cunningham & Werker, 2012):

$$INFO(x,y) = \sum_{i \in X} \sum_{j \in Y/i} p(x_i, y_i)\, log[p(x_i, y_i) / (p(x_i)p(y_i)\,]$$

**Equation 9 Mutual information measure**

$p(x)$ and $p(y)$ are the individual technological profiles of both firms -non-relational information-, and $p(x,y)$ is their mutual technological profile -relational information-. Then, the mutual information of the research or technology profile of the organizations $x$ and $y$ is calculated by considering each particular category $i$ from a larger set of relevant categories (e.g patent technological classes).

- **Min- Complement:** Bar & Leiponen (2012) define the Min-Complement distance as:

$$M(P_i, P_j) = 1 - \sum_{k=1}^{n} min\{p_{ik}, p_{jk}\}$$

**Equation 10 Min-complement measure**

Where $P_i=\{p_{i1},\ldots,p_{ik},\ldots,p_{in}\}$ and $P_j=\{p_{j1},\ldots,p_{jk},\ldots,p_{jn}\}$ are technological patent vectors in $S$ from firms $i$ and $j$. And $S=\{P=(p_1,\ldots,p_n) \mid p_k \geq 0 \; and \; \Sigma^n_{k=1} p_k=1\}$. The Min-Complement distance measure takes values M ∈ [0,1] with M=0 being the closest distance; and satisfies the independence of irrelevant patent classes' property (e.g. a focal firm's distance from other firms only depends on relevant patent technological classes).

- **Mahalanobis:** This measure defines *ex ante* a matrix with the similarity between technologies rather than firms. This similarity matrix is estimated by the technology co-occurrence at the organizational level (Breschi et al., 2003). Then, the similarity of technology is established by the presence of firms in multiple technologies (e.g. survivorship method). Following Stellner (2014) and Bloom et al. (2013), be $F(NxT)$ a matrix with rows $fn$. Then, $\tilde{F}(NxT)$ is a matrix defined as:

$$\check{F} = \begin{pmatrix} f_1 / \sqrt{f_1 f_1'} \\ \cdots \\ f_N / \sqrt{f_N f_N'} \end{pmatrix}$$

**Equation 11 Similarity matrix**

Where $\check{F}\check{F}'$ is the angular separation used by Jaffe (1986). Let $f(:,t)$ be the $tth$ column of $F$. Now, define $\tilde{X}(NxT)$ as:

$$\tilde{X} = \begin{pmatrix} f_{(:,1)} \Big/ \sqrt{\left(f_{(:,1)} f'_{(:,1)}\right)} & \cdots & f_{(:,T)} \Big/ \sqrt{\left(f_{(:,T)} f'_{(:,T)}\right)} \end{pmatrix}$$

**Equation 12 Columns vector of similarity matrix**

Lets define $\emptyset$=$\tilde{X}'\tilde{X}$, which is the angular separation between technologies. $\emptyset$ goes from 1 (similar fields) to 0 (dissimilar fields). Finally, the Mahalanobis distance matrix is define as:

$$M = \tilde{F}'\emptyset\tilde{F}$$

**Equation 13 Mahalanobis distance matrix measure**

Where element $[i,j]$ of this matrix is the distance between firms $i,j$.

- **Weighted angular separation:** Stellner (2014) proposes a measure of similarity between technological fields as in the Mahalanobis distance. The method uses the co-occurrences of patent classification in patent documents to obtain its angular separation (Breschi et al., 2003). In this method the technology field similarity is always the same and is calculated based on the set of patents used in the sample. Let $M$ denote a $TxT$ *ex ante* matrix of technology field similarity. The weighted angular separation is defined as:

$$WA_{ij} = \frac{f_i M f_j'}{\sqrt{(f_i M f_i') * (f_j M f_j')}}$$

**Equation 14 Weighted angular separation measure**

Where $i$ and $j$ are the comparing firms.

- **Aggregated patent-to-patent angular separation:** this measure computes the angular separation from the profile of one particular patent of firm $i$ ($k_{ix}$) and one patent of firm $j$ ($k_{ix}$) and takes the average of all possible combinations of patents (Stellner, 2014). This measure of distance is defined as:

$$PP_{ij} = \frac{1}{X}\frac{1}{Y}\sum_{x=1}^{X}\sum_{y=1}^{Y}\frac{k_{ix}k_{jy}'}{\sqrt{(k_{ix}k_{ix}') * (k_{jy}k_{jy}')}}$$

**Equation 15 Patent to patent angular separation measure**

Where $X$ denotes the number of patents owned by firm $i$, and $Y$ denote the number of patents owned by firm $j$.

## 3.4 SURVEY AND INTERVIEW

In these methods, authors operationalize the technology distance or relatedness concept by scoring in scales from the questions or statements in the interview or survey (Bierly et al., 2009). This kind of methods have allowed to go further in terms of differentiate between factual knowledge and tacit knowledge (Weber & Weber, 2007), operational expertise (De Clercq & Sapienza, 2005), market similarity (Cassiman, Colombo, Garrone, & Veugelers, 2005), and even production similarity (Sapienza et al., 2004) in order to obtain information about how different are the general knowledge bases in the partnership.

There are not major differences in the use of these methods, except by the questions used in the studies. For example, De Clercq & Sapienza (2005) asked to the respondents to indicate in a 5-point scale the degree of similarity in technological knowledge. These items were adapted from the prior research of (Teece, 1986, 1992). For example, the survey included the following questions: "*The CEO and I have worked in very similar functional areas.*", "*Overall, our backgrounds are very different.*" (reverse scored), "*We have different areas of industry expertise.*" (reverse scored), and "*Our experience is based on very similar technological areas.*"

Sapienza et al. (2004) developed a new response scale instead of using prior research. This scale captures actual knowledge relatedness of firms as perceived by the managers. The study asked the respondents whether a knowledge source was unique and specific to each business unit or common and applicable to multiple units. The scale used was: (1 = unique in all or almost all of the business units, 2 = unique in a majority of the business units, 3 = unique in about half of the business units, common across the other half, 4 = common across a majority of the business units, and 5 = common across all or almost all of the business units).

Finally, Cassiman et al. (2005) used a more elaborated survey in order to obtain information about similarity between partners. In this case, firms are classified as having overlapping technology if, before the deal, they respond having R&D projects in the same technological fields and have developed capabilities in the same stages of the R&D projects (pp.205). Additionally, the questionnaire included a section of technology-related elements such as economies of scale and scope in R&D, R&D risk spreading, access to technological resources, reduction of spillovers, and reduction of competition in technology markets.

# 4. TECHNOLOGICAL DISTANCE, ALLIANCES, AND INNOVATION PERFORMANCE

In this chapter the core themes that have been studied in the literature in the context of inter-organizational alliances, technological distance, and innovation are reviewed. Throughout the sections, the relevant research works and findings applicable to the study are appraised. Upon reviewing the various papers it became apparent that they could be categorized into three important parts where the technology distance has a critical role: first, the technological distance and the use of other kind of distances studied to collaborate and innovate; second, the role of technological distance, spillovers, and knowledge flows; and finally, the technological distance and the technology sourcing modes. These three central subjects of the literature on technological distance will be reviewed in their respective order.

## 4.1 TECHNOLOGICAL AND OTHER TYPE OF DISTANCES

A key element in the literature to innovate in alliances is the distance between partner's knowledge. However, besides the technological side there are other categories that play a vital role in alliances contexts. Consequently, the literature extends the concept of distance to include not only the technological but other classes in order to explain the effects in the collaboration and innovation performance. Furthermore, inter-organizational collaboration literature has studied the relevance, ambiguity and relationship among the different kinds of distance dimensions (Knoben & Oerlemans, 2006).

The geographical distance is found side to side to the technological distance in several of the reviewed studies. These two classes of distances are argued to impact directly the knowledge creation output (Greunz, 2003). An interesting relationship has been shown between these two distances: in general, firms choose their partners taking into account their geographical closeness and their technological similarity (Broekel & Boschma, 2012). However, the technological distance, when compared to the geographical distance, shows the most significance using standardize coefficients in the equations (Cunningham & Werker, 2012). Therefore, a suitable choice of technological distance between partners can change a collaboration from a low to a highly productive level. This implies that, although geographical effects are important determinants to R&D collaboration and innovation, these key elements occur most often between organizations not too far from each other in technological space (Scherngell & Barber, 2009).

The importance of these two kinds of distances is shared in different contexts. In university-industry relationships, technological similarity and geographical distance are related to the achievement of higher innovative outcomes (Petruzzelli, 2011). In technology districts, geographical distance and cognitive proximity have a positive relationship as a means for reaching external knowledge sources (Petruzzelli et al., 2009). In industrial R&D compared to public R&D collaboration, geographical factors significantly affect the former, while the effects are smaller in the latter (Scherngell & Barber, 2011). These results are shown also in the mutual dependency between head offices location and the location of knowledge intensive business services, where geographical proximity in itself is neither a sufficient nor a necessary condition, highlighting the idea that other types of proximity such as the cognitive or technological one also play a critical role (Aslesen & Jakobsen, 2007).

In general, technological distance is a stronger factor than geographical distance for cooperative activities (Scherngell & Barber, 2009). On the one hand, cognitive proximity can provide the foundation to collaboration while geographical proximity acts as a reinforcing dimension (Mattes, 2012). On the other hand, technological distance can contribute to bridging geographical boundaries because of the difficulty of overcoming both of them at the same time (Phene, Fladmoe-Lindquist, & Marsh, 2006).

Other types of distances are also found in the literature review escorting the technological distance although they are less representative than their geographical equivalent; these are organizational, social and institutional distances. These distances are also used in some studies about the technological dimension to improve the understanding about collaboration, learning, and innovation performance in the inter-organizational domain. Therefore, it's relevant to define each one of them. Organizational proximity is defined as the distance (closeness) of managerial arrangements about incentives structures and work organization (Boschma, 2005; Meister & Werker, 2004). Social proximity refers to the social embeddedness in terms of friendship, kinship, and experience among actors (Boschma, 2005). Finally, institutional proximity is a normative dimension that refers to institutional properties: laws, rules, norms, values and routines that form the socio cultural, economic, and political framework in which the actors are embedded (Boschma, 2005).

In general, the effect of organizational proximity in innovation and collaboration is not conclusive. On the one hand, some studies argue that in contrast to the technological dimension, the organizational proximity influences indirectly the output of collaborations (Cunningham & Werker, 2012). On the other hand, other studies establish that similar to the organizational distance but from a knowledge base content specific component, the managerial knowledge distance has an effect as much as critical than the technological distance in the collaboration output among firms (Schulze & Brojerdi, 2012). Furthermore, organizational proximity seems to strengthen the connection and exchange of knowledge between organizations while having an opposite influence or no influence on the innovative performance of firms (Broekel & Boschma, 2012). As a consequence, the role of the organizational proximity is normally set to support the bases for collaboration (Mattes, 2012) and to be used as an approach for reaching external knowledge sources (Petruzzelli et al., 2009).

Similarly to technological proximity, social proximity increases the likelihood to connect and exchange knowledge (Broekel & Boschma, 2012). Moreover, the social dimension plays a critical role in the results of knowledge intensive business services and clients relations (Aslesen & Jakobsen, 2007) and is useful to support informal communication (Mattes, 2012). Nevertheless, in contrast to organizational proximity which main function is to create and sustain linkages, the social dimension also facilitates nurturing innovation performance (Broekel & Boschma, 2012). Institutional proximity is the less mentioned dimension in the review. However, it has been noticed that in synthetic knowledge bases the institutional dimension and its cognitive counterpart are the most important types of proximities. Furthermore, the institutional proximity, like the organizational proximity, provides foundations for collaboration among organizations (Mattes, 2012).

## 4.2 TECHNOLOGICAL DISTANCE, SPILLOVERS AND KNOWLEDGE FLOWS

In general, foreign knowledge can influence endogenous knowledge in two ways. The first effect is that foreign technologies might be simply transferred such as introduced in foreign markets and adopted. The second one, is related to the flow of knowledge across borders making that foreign inventions increase productivity through endogenous innovation (Pizer & Popp, 2008). These two forms of how external knowledge influences internal knowledge have been named as embodied rent spillover and disembodied knowledge spillovers respectively (Griliches, 1979, 1992). On one hand, rent or market spillovers (Jaffe, 1986) arise if goods are not priced at their user value due to quality enhancements that are not reflected in the pricing of the good (Griliches, 1979). On the other hand, technology knowledge spillovers are defined as the non-appropriable amount of technological knowledge that is produced by a firm's innovation effort (Kaiser, 2002) and as a result they depend on the technological edge that the external firm has over the domestic firm (Peri & Urban, 2006). The focus of this section is on the second form, where the increase in productivity is the result of knowledge '*spillovers*' (Verdolini & Galeotti, 2011).

Knowledge spillovers are driving elements in the innovation process. Innovation contributes to the creation of new knowledge which diffuses over technological distances (Greunz, 2003). As a result, some studies denote the disembodied knowledge spillovers by using the concept of technological distance (Branstetter, 2001; Jaffe, 1986) probably because among the various channels used by this kind of spillovers (e.g. patents, foreign direct investment, or the presence of foreign firms) one of the most important is the technological similarity (J. W. Kim & Lee, 2004). Moreover, measuring spillovers is a difficult task and therefore the use of proxies such as proximity or technological distance between actors is common (Kaiser, 2002); for example, based on patents which are used to generate appropriability.

The higher the distance between actors in technological space, the lower the probabilities of knowledge flow between them (Verdolini & Galeotti, 2011). Technological knowledge diffusion is enhanced by physical and technological proximity, which means that the inventions are similarly distributed across technical fields (MacGarvie, 2005). For example, in a study of knowledge diffusion using patent citations[3] as an indicator of knowledge flows, it was found that citations to

[3] Citations are used as indicators of knowledge spillovers based on the argument that they embody a connection to the pre-existing knowledge upon which the invention is built (Criscuolo & Verspagen, 2008).

patents in the same patent class as the citing patent are over 500 times as likely as citations to patents in other classes (A. G. Z. Hu & Jaffe, 2003).

In general, spillovers have a substantial positive influence on innovation and technology. Domestic R&D and the disembodied spillovers proxied by the technological distance affect the Total Factor Productivity changes through technical progress (J. W. Kim & Lee, 2004). The presence of foreign firms benefit local firm's growth[4] and productivity not only by the presence of foreign direct investment, but by the effects of the technological edge of the foreign firm joined to the technological proximity of the local firm (Peri & Urban, 2006). Even when technological distance is large, the use of mechanisms associated to inter-firm knowledge flows such as mobility and alliances help to counteract its negative effect (Rosenkopf & Almeida, 2003). Moreover, the spillovers from technological proximate R&D stocks are both significant and important at different contexts such as intra industrial agglomerations (Fung & Chow, 2002; Orlando, 2004), cross-industry innovation (Enkel & Gassmann, 2010), sub-national regions (Greunz, 2003), and R&D cooperation networks (Cantner & Graf, 2006).

## 4.3 TECHNOLOGICAL DISTANCE AND TECHNOLOGY SOURCING MODES

Organizational boundary spanning can be achieved by sourcing external technology. Different external technology sourcing modes can be found in different contexts ranging from individual firms to inter-organizational relationships. On the one hand, technological distance can be used when the analysis is centered in a focal firm. For example, in innovation search theory, technological distance is used to understand how far from the existing technological portfolio in-licensing firms are able to move when they in-license externally developed technologies (Laursen et al., 2010). Also, to understand how knowledge-similarity is a key factor in affecting

[4] This channel of spillovers is defined as the Veblen–Gerschenkron effect of foreign direct investment.

firms' technological diversification (Breschi et al., 2003); or in how portfolio technology relatedness moderates the relationship between technological diversification and technological performance (Leten, Belderbos, & Van Looy, 2007). In the new product development context, technological distance is used to find why firms that are located near the industrial technological frontier have more probabilities to survive than those firms that are far from it (Fontana & Nesta, 2009). On the other hand and aligned to this section, technology distance is used in contexts in which the firms are engaged in inter-organizational relationships to create new knowledge or technologies. The most common inter-organizational relationships are operationalized through strategic alliances, merger and acquisitions, and other kind of sourcing modes such as corporate venture capitalist (van de Vrande et al., 2011).

In the first place, it's generally accepted that alliances are a critical mechanism to create new technological knowledge (van de Vrande et al., 2011). Empirical results have shown that technological distance and innovation performance in alliances are related in a U-inverted curvilinear way (C. Kim & Song, 2007; C. Lin et al., 2012). This inverted U-shaped effect of technological distance remains over time (Schildt et al., 2012), in contexts such as in university-industry collaborations (Petruzzelli, 2011), and is known as the proximity paradox (Broekel & Boschma, 2012). However, the positive effect for firms is much higher at the long term of the relationship and when engaging in more radical, exploratory than exploitative alliances (Bierly et al., 2009; Nooteboom et al., 2007). Consequently, the idea that some but not too much technological distance between collaborating actors increase innovation outcome (Fornahl et al., 2011) has resulted in the search of an optimal technological distance (Wuyts et al., 2005). Such relation suggests that not all strategic alliances have to be treated as equivalent; depending on the technological distance, some firms are more convenient than others in order to increase the innovative performance. Finally, the inverted U shaped and optimal distance are also critical elements at the network alliance level in the knowledge exchange and innovativeness of the embedded firms (Boschma & Frenken, 2010). The network firms should focus in the quality of linkages into the network by selecting those partners having the optimal technological distance in order to increase learning and innovation (Fornahl et al., 2011). Therefore, average technological distances or intermediate amounts of similarity between partners have been elements used even to explain network formation (Cowan & Jonard, 2008) and the advantage position of companies in the network to develop innovations (Gilsing et al., 2008).

In the second place, studies about the innovation of mergers and acquisitions (M&A) have concentrated on the role of technological knowledge. Technological proximity has been used to detect and assess companies to support M&A target choice decision making (H. Park, Yoon, & Kim, 2013). However, the influence of M&A is in general adverse in the creation of innovative technologies; this result is affected by the technological relatedness of technological portfolios of the partnering firms, since larger technological distances strengthen the negative effect in M&A (van de Vrande et al., 2011). This kind of results contradict the idea that higher levels of technological relatedness between merging parties are associated with better post-merger outcomes (Ornaghi, 2009). Furthermore, studies have shown that the R&D inputs are reduced after the M&A when merged entities are technologically similar (Cassiman et al., 2005). Consequently, knowledge similarities in M&A facilitate incremental renewal, but had no effect on invention quantity or quality, and has a negative effect on invention novelty (Makri et al., 2010).

Finally, technological distance has an important role on other types of relationship sourcing modes. For some studies, technological distance does not play a role in the interaction with corporate venture capital (CVC) investments on the creation of pioneering technologies (van de Vrande et al., 2011). However, knowledge relatedness has demonstrated to facilitate knowledge transfer and creation between CVC firms and their innovative portfolio companies (Weber & Weber, 2007). Moreover, has been shown that lower levels of knowledge overlap between the CVC and their portfolio firms were related with superior learning in a linear fashion (De Clercq & Sapienza, 2005). In vertical sourcing modes such as the value chain formed by innovation firms and their suppliers, the technological distance operationalized through the differences of inter-industry knowledge bases has shown a positive effect on the likelihood of generating pioneering innovations (Li & Vanhaverbeke, 2009). Consequently, knowledge overlap between suppliers and customers are central for projects concerning new technologies (J. Lee & Veloso, 2008). This value chain relationship has shown an inverted U-shaped of cognitive distance effects on innovation performance –similar to the alliances behavior- (Bönte & Wiethaus, 2007). In addition, technological distance is highlighted as a positive key element in other external knowledge sourcing mechanisms such as firm participation in technical committees (Nambisan, 2013), relationships between head offices and knowledge intensive firms (Aslesen & Jakobsen, 2007), leaders and followers firms in product market competition (Aghion, Bloom, Blundell, Griffith, & Howitt, 2005), and spin-off firms and its partners (Sapienza et al., 2004).

# 5. HYPOTHESES

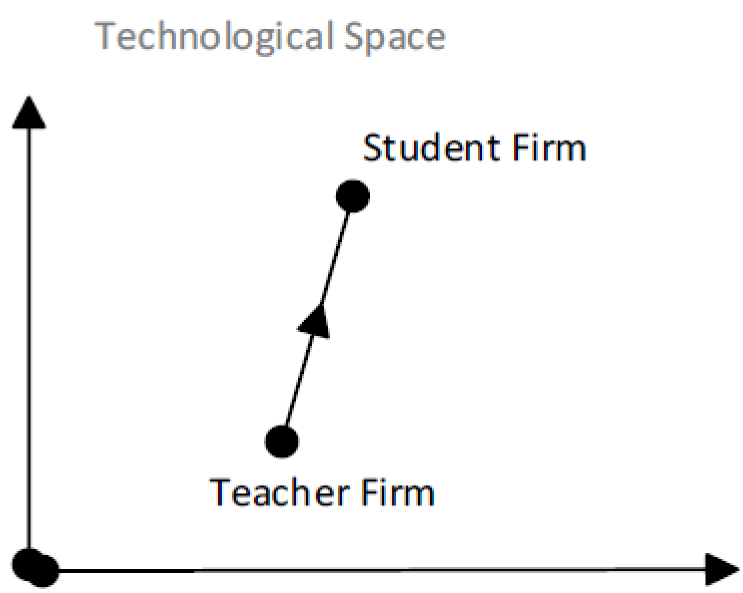
This section presents the main hypotheses of the study. It is contended that the literature's notion of technological distance between two firms in an alliance can be broken down into two concepts, student firm technological distance and teacher firm technological distance, with respect to their common technological innovation, and claim that these different types of technological distances have an effect in the innovation performance -specifically in the technological value of their join innovation- for the organizations. Also, a claim is made in terms of the effect of technological capital of individual alliance's firms on the value of joint innovation under the framework of learning dyad. In developing the hypotheses, the present thesis builds upon the knowledge -based view and the relative absorptive capacity perspective.

## 5.1 BACKGROUND

For clarity, the following definitions are used. First, following Lane & Lubatkin (1998), this study distinguishes between two types of organizations in an alliance: student firm and teacher firm. The notion of the pair student-teacher firms comes from what the authors called the '*learning dyad*' in an alliance which is an extension of the concept of absorptive capacity (W. M. Cohen & Levinthal, 1990) at

an inter-organizational level and the flow of knowledge between both of them in the learning process. Then, in the alliance "*the ability of a firm to learn from another firm is jointly determined by the relative characteristics of the student firm and the teacher firm*" (Lane & Lubatkin, 1998, pp.462). These relative characteristics, such as the distance of knowledge bases in technological space, are found between each one of the firms in the alliance. Figure 1 shows the location of firms, black dots, and their technological distance in a technological space of two dimensions without loss of generality. Although the argumentation is stated in a one-way inter-organizational learning, the factors that influence one-way learning, also have an effect in two-way learning. Therefore, the study represents the contingencies that influence absorptive capacity in all inter-organizational learning.

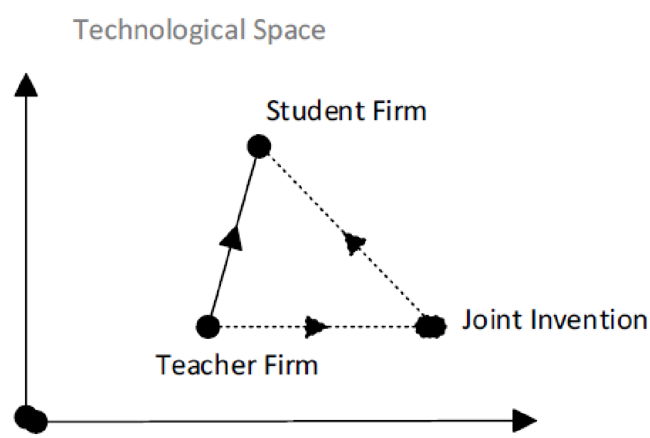
**Figure 1. Technological distance between firms**




Second, the two kinds of organizations lead to two types of what this study calls '*relative technological distances*', the student firm technological distance and the teacher firm technological distance. To understand this argumentation is important to develop some ideas further. It is assumed that firms in an inter-organizational relationship, such as alliances, generate new knowledge which can be represented in new joint technological knowledge or joint inventions. This joint technological invention has its own knowledge base and can be located in the technological space. The location of the joint invention in the technological space originates two technological distances to each one of the firms involved in the alliance. This study calls these distances '*relative technological distances*' because they are relative to the joint invention produced by the two firms. A firm from the learning dyad can be

further away from its join invention than its counterpart in the alliance. This study defines, by comparing between partners, the student firm as the firm with the narrower scope of knowledge and expertise in the learning dyad, and the teacher firm as the firm with the broader scope of knowledge and expertise in the learning dyad. Therefore, this study assumes that in the technological knowledge domain in an alliance, the firm with the more knowledge and experience '*teaches*' to a learning '*student*' firm which has less knowledge and general expertise. Figure 2 shows the location of firms, their joint invention, and their respective relative technological distances, as dashed lines, in a technological space of two dimensions

**Figure 2. Relative technological distances**




Third, the intellectual capital has a fundamental role in modern business organizations and has been identified as a critical driver of firm performance (Teece, 1998; Youndt, Subramaniam, & Snell, 2004). In general, intellectual capital has three categories: human capital, relational capital, and structural capital (Edvinsson & Sullivan, 1996). From those three categories, structural capital is referred to the processes and procedures created by, and store in, firms technology systems that speed the knowledge flow through the organization (Carson, Ranzjin, Winefield, & Marsden, 2004; Youndt et al., 2004). Moreover, structural capital is also related to the knowledge that has been captured by the firm and embedded in the organization (Jansen, Tempelaar, van den Bosch, & Volberda, 2009). Since intellectual capital is the ability to translate new ideas into products or services (Booth, 1998), or in other words, to convert invisible assets such as knowledge into resources (Bradley, 1997), then knowledge can be converted into something owned by the firm such as technology (e.g. patents) (Hsu & Wang, 2012). As a result, this technological capital is related to the capacity to invent new technology and to innovate or commercialize that technology (Flavio & Avila, 2010). Taking into account that intellectual capital is seen as knowledge that can be converted into profit (Harrison & Sullivan, 2000), or better, into value (Edvinsson & Sullivan, 1996); and that with a strong structural capital, firms value creation activities will be more efficient and effective (Bontis, 1998; Widener, 2006); then, technological capital as an inner type of intellectual capital is related to the innovation performance and value creation of firms (Han & Li, 2015).

In the following section, this study associates the hypothesis to the two relative technological distances and the firms' technological capital in the alliance.

## 5.2 HYPOTHESIS 1

Research studies suggest that the heterogeneity of resources have concentrated on factors like the origin, the access and the maneuvers used by actors in different contexts (Ahuja & Katila, 2004; Quintana-Garcia & Benavides-Velasco, 2006; Rosenkopf & Almeida, 2003; Rosenkopf & Nerkar, 2001). Some studies have focused on the implications of resource heterogeneity on the innovation performance (Jiang, Tao, & Santoro, 2010; Nooteboom et al., 2007; Sampson, 2007). Technological alliances are means trough which firms find, access, and combine a variety of resources in order to generate innovations and competitive advantages (Ahuja & Katila, 2004; Porter, 1990; Rosenkopf & Almeida, 2003). Among the resources, knowledge is the primary resource underlying the creation of value and competitive advantage (Barney, 1991; Grant, 1996; Kogut & Zander, 1992).

In alliances, firms identify, assimilate, and exploit knowledge from each other to jointly generate joint new knowledge (Lane & Lubatkin, 1998). In this process, firms re-combine their knowledge-base resources (Greeven & Xiaodong, 2009) and increase their innovation capability. This results in a spanning of their technological boundaries (Rosenkopf & Nerkar, 2001) by exploring new external knowledge embedded in their innovations (Andriopoulos & Lewis, 2009). Therefore, the external knowledge that goes beyond the firms' technological fields leads to innovation and their technological expansion (Rosenkopf & Nerkar, 2001). In an alliance, the exposure to technological knowledge improves each firms' ability to apply this knowledge and then improve their innovation quality (Ahuja & Lampert, 2001). This could be attributed to the addition of novel and valuable resources into a firm's innovation system originated from the technological similarity or dissimilarity to the external knowledge (Phene et al., 2006). Thus, this technological distance from the firm to the external knowledge influences its resulting innovation quality (W. L. Lee, Chiang, Wu, & Liu, 2012). Since past studies, the technological quality of an innovation has been related to its technological value (Albert, Avery, Narin, & McAllister, 1991; Green & Scotchmer, 1995). Furthermore, the higher technological quality of the invention, the more inventions should build upon the underlying innovation increasing its value (Fischer & Leidinger, 2014). Therefore, the technological distance from the firm to the external knowledge is likely to affect the value of the firm's invention. This argument can be applied to each one of the learning dyad firms in the alliance (student and teacher firms) where the technological distances are relative to their joint invention. Based on the above arguments the following base hypothesis is suggested:

**H1a.** *Relative technological distances have an effect on the value of the joint invention.*

It is generally accepted that inter-organizational collaborations, such as alliances, are used to share resources and increase innovation performance (Pitsis & Gudergan, 2010). However, alliances also have high uncertainty, complexity, risk, and differences in the goal among the partner firms (Das & Rahman, 2001) that make them often to fail (S. H. Park & Ungson, 2001). Therefore, taking into account that learning is an essential condition for the efficiency of innovation activities and the need to decrease the uncertainty generated, it´s likely that the introduction of innovations might not take place too far away from the competences of the firm (Antonelli, 2004). Therefore, firms search in close proximity to their existing knowledge base to reduce uncertainty and increase innovation performance (Boschma, 2005). This is in line to other studies that have found that in general alliances specialize in a certain research field, rather than to enter a completely new market (Zidorn & Wagner, 2013). Because in the learning dyad both firms will share the same potential jointly created knowledge, it is probable that the new created knowledge is closer to the firm’s respective knowledge bases, rather than far from it, to reduce the uncertainty generated resulting from the inter-organizational collaboration and to seek opportunities from the other firm’s knowledge base (Colombo, 2003). Then, it is expected that, in the learning dyad, the relative technological proximity have a positive effect on innovation value. Or, in other words, that both relative technological distances of firms in the learning dyad have a negative effect on innovation value. Hence, it is hypothesized that:

**H1b.** *The higher the relative technological distances of both firms in the learning dyad, the lower the value of the joint invention.*

## 5.3 HYPOTHESIS 2

Learning is an implicit objective in all alliances (Kale & Singh, 2000; Yoshino & Rangan, 1995). The distinction between the student and the teacher firm in a learning dyad alliance and their relative technological distances is based on the approach of the relative absorptive capacity concept (Lane & Lubatkin, 1998). Applying the concept of absorptive capacity to an alliance suggests that acquiring external knowledge to innovate can be seen as a learning platform (C. Lin et al., 2012). In the present study, the teacher firm is defined as having broader knowledge and expertise than the student firm. This brings an important argument to be discussed. The creation or addition of joint knowledge and competences impacts differently the teacher and student firms in terms of how familiar or unfamiliar is the new joint developed knowledge to their respective knowledge bases. Because the teacher firm has a broader knowledge than the student firm, then it is likely that the new knowledge is more related to its knowledge base than the student's. As Cohen & Levinthal (1990) pointed out in their central argument about absorptive capacity: *"… a diverse background provides a more robust basis for learning because it increases the prospect that incoming information will relate to what is already known.*"(pp.131). Therefore, the jointly created or added knowledge or competence makes the teacher firm to have a more exploitative learning, when compared to the student firm, and is expected to change in a lesser grade the basic nature of its innovative activities (Hagedoorn & Duysters, 2002; Rowley, Behrens, & Krackhardt, 2000).

On the other hand, the student firm can be characterized by having a greater disturbance for the new jointly created or added knowledge when compared to the teacher firm. For the student firm, the new jointly created knowledge should go further from its knowledge base than the teacher firm, having a more explorative learning (Almeida & Kogut, 1999; Fleming, 2001; Rosenkopf & Almeida, 2003; Rosenkopf & Nerkar, 2001). This balance of exploration and exploitation (March, 1991) between the two firms is found in strategic alliance formations (Z. Lin, Yang, & Demirkan, 2007; Rothaermel & Deeds, 2004). Furthermore, this simultaneous acting of innovation ambidexterity of exploration and exploitation (M. H. Lubatkin, Simsek, Ling, & Veiga, 2006; Simsek, 2009) makes more likely to achieve better organizational performance (Junni, Sarala, Taras, & Tarba, 2013). Succeeding ambidexterity produces outcomes that are not achievable if one of these dimensions, exploration or exploitation, is accentuated at the cost of the other (Cao, Gedajlovic, & Zhang, 2009; Gibson & Birkinshaw, 2004), because they have a complementary and reinforcing effect on performance (Gupta, Smith, & Shalley, 2006). Then, in the learning dyad the symbiotic relationship of exploration and exploitation (Garcia, Calantone, & Levine, 2003) can be seen reflected in student and teacher firms, respectively. The increasing relative technological distance might potentially turn the exploitative dimension of ambidexterity of the teacher firm into a more explorative profile. However, increasing the relative technological distance will keep the explorative dimension of the student firm. In high relative technological distances, both firms in the learning dyad can be considered as explorative and the innovation ambidexterity is lost resulting in lower organizational performance. Therefore, and reminding the negative effect of the relative technological distances on join invention value, it is hypothesized:

**H2.** *In the learning dyad, the relative technological distance of the teacher firm has a larger negative effect on the value of joint invention than the relative technological distance of the student firm.*

## 5.4 HYPOTHESIS 3

The absorptive capacity of knowledge depends of the cumulative knowledge of firms. Organizations need a stock of prior knowledge to assimilate and use new knowledge (W. M. Cohen & Levinthal, 1990). The absorptive capacity is a learning and problem solving capability that helps in the assimilation and creation of knowledge (Mowery & Oxley, 1995). The dependence on cumulative knowledge of absorptive capacity has made it equal to the knowledge bases of firms (Ahuja & Katila, 2001; L. Kim, 1998; Mowery et al., 1998). Consequently, absorptive capacity more often has been operationalized as the R&D intensity used to build that knowledge base (Meeus, Oerlemans, & Hage, 2001; Mowery et al., 1996; W. Tsai, 2001). Technological knowledge is revealed in patents; thus, patents can also be a reflection of absorptive capacity (Ahuja & Katila, 2001; Mowery et al., 1996). Patents are an exemplification of the technology capital belonging to the structural category of the intellectual capital of firms (Hsu & Wang, 2012). The absorptive capacity of firms is a critical issue of a firm's ability to create new knowledge (W. M. Cohen & Levinthal, 1989). Therefore, it's expected that technological capital is a critical element in a firm's innovation performance.

Accordingly, firms with larger technological capital (e.g. patents) have more absorptive capacity and can deal with technological distance in a better way than firms with smaller technological capital. To further study this, it is differentiated between the student technological capital and the teacher technological capital contexts. Firms can conceive a technology patent stock which is a result of their effort in R&D. Then, this portfolio or technological capital is a representation of the technological codified knowledge that has been created. Technological capital helps in the absorption of knowledge that is embedded in an innovation located from a technological distance from the firm.

Given the characteristics of the student firm, which is more focused on exploration and novelty when compared to the teacher firm, large amounts of technological capital reinforce the ability of the student firm to deal with its larger relative technological distance to the join invention. This has a positive effect on the innovation value. By contrast, given the characteristics of the teacher firm, which is more focused on exploitation and incremental improvements when compared to the student firm, it's not expected that technological capital be as important as it is for the student firm. Then, a limited effect of technological capital on the ability of the teacher firm, when compared to the student firm is expected. Therefore, it is hypothesized that:

**H3.** *In the alliance, the student's firm technological capital has a positive and higher effect on the value of the joint invention than the teacher's firm technological capital.*

# 6. RESEARCH STRATEGY AND METHODOLOGY

To address the hypotheses and thereby implement this study, a positivist paradigm strategy and quantitative methods are used. To identify the external reality, the positivist researcher eliminate alternative explanations and allow fundamental factors to be measured accurately in order to test preset hypotheses (Easterby-Smith, Thorpe, et al., 2008). The aim of this section is to expose the general philosophical assumptions that underline this research (Section 7.1), as well as to present the methodological approach under consideration (section 7.2)

## 6.1 RESEARCH APPROACH

The lack of philosophical thinking, although is not fatal, has a serious impact in the quality of research (Easterby-Smith, Thorpe, et al., 2008). Researchers have a '*worldview*' (Creswell, 2014) based on assumptions about how and what they learn and a set of beliefs that guide their actions (Guba, 1990). This '*worldview*' has been also called '*paradigm*' (Denzin, Lincoln, & Guba, 2011) which informs and guides the researcher with regards to epistemology, ontology and methodology. Therefore, the paradigm positioning was the first step in defining the research approach of this study.

The present research subscribes to the positivist paradigm. The positivist movement emerged with the work of August Compte (Comte, 1896) who argued that if much of the physical world operates according to laws, so does society (Macionis, 2012). Simply explained, positivism assumes an ontological realism where the world or reality is concrete and external; and an epistemological objectivism where knowledge and science progress is based on the objective observation of this external reality (Denzin & Lincoln, 2005).

To identify the external reality, the positivist researcher eliminates alternative explanations and allows fundamental factors to be measured accurately in order to test preset hypotheses (Easterby-Smith, Thorpe, et al., 2008). Thus, positivism holds assumptions more for quantitative research, where the development of numeric measures of observation are dominant (Creswell, 2014). Moreover, progress in research by positivism is made through hypothesis and deductions where ideas are not often induced (Easterby-Smith, Thorpe, et al., 2008). Consequently, the methodology used in this study has a quantitative and deductive approach by first developing hypotheses and then testing the results of empirical observation and measurement (Lancaster, 2005). Table 1, shows concisely the methodological implications of positivism and contrasts them to its counterpart relativism and constructionism.

The followed deductive research process, which in general is called the scientific method (Creswell, 2014), formulates objective theories or hypotheses that are operationalized into indicators or measures (e.g variables representing the reality), and then the resulting data is analyzed using statistical processes in order to test the initial theory or hypothesis. Specifically, the present study makes use of hierarchical or sequential regression methods as a kind of multiple regression in order to examine the specific scientific hypotheses and relationships among data.

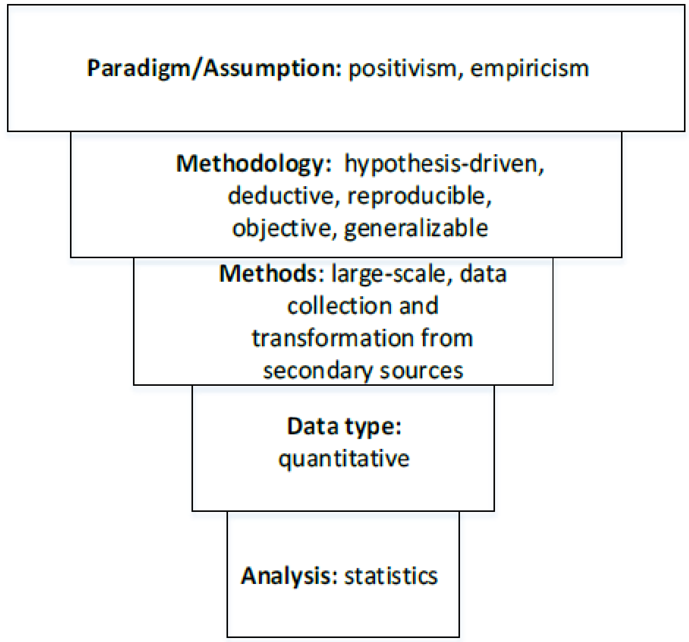
**Table 1. Methodological implications of different epistemologies**

| Methodological elements | Positivism | Relativism | Constructionism |
|---|---|---|---|
| Aims | Discovery | Exposure | Invention |
| Starting points | Hypothesis | Propositions | Meanings |
| Design | Experiment | Triangulation | Reflexivity |
| Techniques | Measurement | Survey | Conversation |
| Analysis/interpretation | Verification/ falsification | Probability | Sense-making |
| Outcomes | Causality | Correlation | Understanding |

Fuente: (Source: Easterby-Smith, Thorpe, et al., 2008, pp.63)

In general, the multiple regression method is a strategy to explain or predict a criterion or dependent variable, with a set of predictor or independent variables (Petrocelli, 2003). The multiple regression methods can be divided in simultaneous, stepwise, and hierarchical regression. The present study uses the hierarchical regression method to involve theoretically based decisions for how predictors are entered into the analysis in order to examine the formulated hypotheses (B. H. Cohen, 2013). The interest of hierarchical or sequential regressions is to test and examine the influence of predictor variables in a sequential way to understand how each predictor changes the prediction. The change in *R2* or the Log likelihood, and the *p* values, are the statistics of interest in hierarchical regression (Wampold & Freund, 1987). This method is widely used in the empirical technology management research on the quantitative approach dimension to analyze the effect of different predictors in the models. Finally, Figure 3 resumes the approach followed by this research and its methodological path.

**Figure 3. Research approach**




## 6.2 METHODS, TOOLS, AND MEASURES

This subchapter describes the research methods used in the present study. The subchapter is divided in three sections. The first section describes the setup, data collection and the sample details. In the second section, the variable definitions are presented. The third section specifies the models used in testing the hypotheses, and describes the statistical methods.

**6.2.1 Set Up, Data Collection, And Sample Details** The present study tests the hypotheses on a data set of 465 joint patent applications involving firms in the biotechnological industry from 2006 to 2010. This section offers information about the time-frame selected, the joint patent applications, the industry selected, the patent database used, the method to identify the patents, and finally the collection process and data sample used in the research.

**6.2.1.1 The time-frame period.** The time frame used in the present research ranges from 2006 to 2010 -five years-. The five year time-frame has been shown to be a good time in research of technological distance, R&D alliances, and innovation performance of firms (C. Lin et al., 2012; H. Park et al., 2013). Moreover, the five year time-frame have been already used in contexts such as join patenting in inter-organizational collaborations (Petruzzelli, 2011) and technological relatedness in patent portfolio analysis (Leten et al., 2007). This time-frame is particularly useful in rapid change high technology industries such as biotechnology, probably because the life span of alliances is usually no more than five years (Gulati, 1995b). Even recent studies have assumed shorter average duration of alliances: three years (Lavie & Miller, 2008). A possible explanation for this argument, as evidence suggests, might be that the technological knowledge that drove alliances in their initial arrangement could be of little or non- significance few years later (Schildt et al., 2012). Therefore, for this research a five year period is a reliable and appropriate time-frame for assessing technological impact (Ahuja, 2000; Henderson & Cockburn, 1996; Podolny & Stuart, 1995; Stuart & Podolny, 1996; Vanhaberbeke, Duysters, & Beerkens, 2002).

**6.2.1.2 Joint patent applications** A joint patent or co-patent is defined as the patent owned by two or more firms (Briggs & Wade, 2014) that form a duopoly or tight oligopoly comparable to a restrictive licensing agreement from an economic standpoint (Aoki & Hu, 1999). Accordingly, the shared property rights of join patents must not be confused from other multi-party agreements such as cross-licenses, licenses for reciprocity, polled patents, and patent infringement agreements (Hagedoorn, 2003). Because its nature, join patenting is generally linked to R&D partnership (Hagedoorn et al., 2003) such as inter-firm (Briggs, 2015) or university-firm collaborations (Petruzzelli, 2011). Although join patenting research is an area minimally studied (Briggs & Wade, 2014), the growth of research in the area is increasing, probably because the overpassing in their higher quality characteristics (Briggs, 2015) when compared to single owned patents. The present research is based on the assumption that if collaborative R&D can be measured by patent indicators, then join application patents should be a good measure of innovative output in an alliance (C. Kim & Song, 2007). Therefore, a joint patent can be assumed as an alliance outcome of the R&D collaborative type. Thus, although the critical aim of the present study is not approaching the gap of joint patent research, it indirectly does it, by including them as a unit of analysis in order to improve the understanding of innovation in alliances or inter-organizational collaboration.

Following Noteboom (2007), the patents used in this research are those that have been successfully applied for. Patent applications have already been used in different studies as an indicator of innovativeness (Wagner & Cockburn, 2010), output of R&D activity (H. Park et al., 2013), to identify pioneering technologies (van de Vrande et al., 2011), in technological diversification (Leten et al., 2007), to analyze networks of innovators (Cantner & Graf, 2006), and in the technological distance domain (Buerger & Cantner, 2011). Patent applications are good indicators of firms' technological competence because when a firm applies for a patent in a technological domain it's assumed that such a firm is at, or close to, the technological frontline and has improved technological capabilities in that domain (Breschi et al., 2003). Therefore, patent applications are even seen as expressions of technical success (Ernst, 2001). However, a critical argument to use patent applications in this research is that patent applications are cited not only by other applications, but also by granted patents which is a core characteristic of the variables involved in this study. Even more, while the patent application is pending, it can be exploited using mechanisms of marking, selling and licensing; it is not necessary to have a granted patent to make economical or technological transfer processes. Therefore, a patent, granted or not, with commercial value or not, is the reflection of R&D efforts and thus gives technological insight that helps to subsequent developments in technology (Bradford & Rajat, 1989; Ernst, 2001).

**6.2.1.3 Biotechnology Industry** As explained earlier, this study focuses on the biotechnology industry. The motives to choose this industry are described in the following rationale. It's generally accepted that biotechnology is a key economic innovation strategy (Fornahl et al., 2011). The biotechnology industry has been recognized as having high alliance collaboration frequency (Luo & Deng, 2009; Powell, Koput, & Smith-Doerr, 1996) particularly because of the rapidly evolving environment and high uncertainty (Pangarkar, 2003). The alliances in the biotechnology industry can include collaborative R&D, licensing agreements and marketing and distribution agreements. The present study emphasizes on the collaborative R&D type, taking into account the potential lead to patentable knowledge that can be represented into join application patents (C. Lin et al., 2012). Patents in the dynamic and knowledge intense biotechnology industry are an indication of innovative success and the generation of new knowledge (Calabrese, Baum, & Silverman, 2000; Newman & Hanna, 2006). Besides, the biotechnology firms' performance depend on their capabilities as reflected in the quality of their patents (DeCarolis & Deeds, 1999). Therefore, the biotechnology industry is appropriate for measuring innovation outcomes such as join application patents because they are important means of intellectual property protection for the firms in this industry (C. Lin et al., 2012). Furthermore, not only the generation of knowledge and its protection is important in the biotechnology industry, but the time to patent became a critical issue where the innovation cycle shortens and quickens making biotechnology a rich patent data industry to analyze (Tzabbar, Aharonson, Amburgey, & Al-Laham, 2008). In conclusion, taking into account that the biotechnology industry is one of the most innovation intensive industries (Phene et al., 2006), it offers an opportunity to examine innovation by firms' collaborative joint technological patents which is the core of the present research.

**6.2.1.4 The USPTO Patent database** Joint patent applications in the biotechnology industry were collected from the applications filled in the United

States Patent and Trade Organization database. The United States Patent and Trademark Office (USPTO) is an agency for granting

patents and registering trademarks[5]. The USPTO advises on intellectual property (IP) policy, protection, and enforcement; and works with other agencies around the world via international agreements. The USPTO has online search of full text and images databases of patents and application patents since 1976 to present. The database has information of patent application and all classes of patents such as utility, design, reissue, plant patents, and includes bibliographic data, full description of the invention, and the claims.

Although the use of this database could imply a bias in favor of companies from United States and against firms from other countries, there are two main reasons to choose it. On the one side, the USPTO database reflects the country with the most dynamic market and R&D efforts (e.g. patents) in the biotechnology industry, higher number of biotechnology companies, and bigger business enterprise expenditures on research and development for biotechnology (OECD, 2013). According to the OECD (2013), the United States -US- is the country with the higher share in biotechnology patents filed under the Patent Cooperation Treaty –PCT- with a 40.94% overpassing the second place that includes the 28 countries of the European Union which together sum up a share of 26.05%. Thus, companies from other countries that look for central positions in their product markets also widely file patents in the US because the importance and the technological sophistication of the US market, and the genuine patent protection provided by the authorities (Hagedoorn, 2003; Patel & Pavitt, 1991). Therefore, even if a biotechnology invention was produced in any country other than the United States, there is a high potential creation of value by filling the invention in the USPTO

[5] In Colombia the agency in charge of these processes is the Superintendencia de Industria y Comercio (SIC).

because it protects it from external use by entering one of the largest patent markets in the world.

On the other hand, although there are other organizations in charge of patenting processes around the world, only one was chosen in order to have consistency, reliability and comparability in terms of the patenting processes because different nations have different patent systems in terms of application, protection and granting (Ahuja, 2000). Regarding prior art, USPTO examiner practices are substantially different from other organizations such as the European Patent Office –EPO- (Alcácer & Gittelman, 2006). For example, in the USPTO the inventor and the lawyer have the responsibility to deliver the whole prior art references considered relevant to the patentability of the invention, something called “duty of candour” (Criscuolo & Verspagen, 2008). In contrast, the EPO does not uses the “duty of candour” (Akers, 2000; Meyer, 2000; Michel & Bettels, 2001) and the patent application should just indicate the background regarded as useful for understanding the invention. Consequently, using different patent systems can have a strong effect in the present study.

**6.2.1.5 Identifying biotechnology patents: using the International Patent Classification System –IPC-** Biotechnology patents were identified using the International Patent Classification (IPC) system. The IPC offers a hierarchical system of symbols to classify patents and utility models in technology areas with the purpose of establishing a search tool to retrieve documents by intellectual property offices and other users (WIPO, 2015). The IPC codes are included in patents and applications as bibliographic data besides typical information such as inventor, country, claims, or assignee which helps to process the information. The IPC was stablished by the Strasbourg Agreement in 1971 administered by the World Intellectual Property Organization –WIPO– that divides technology in eight main sections and about 70,000 subdivisions.

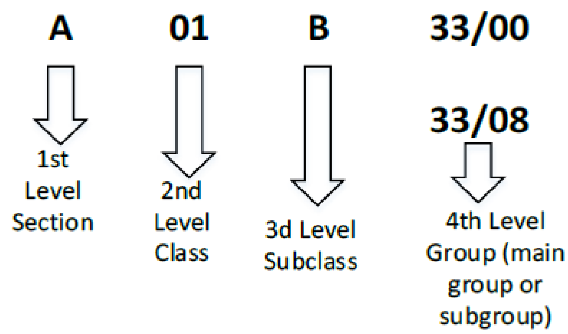
The main eight sections are the highest level of hierarchy of the classification and is designated by one capital letter from A to H[6]. The main sections are subdivided in classes which are the second hierarchical level of the classification, and consists of the section symbol followed by a two digit number (e.g. H01 Basic electric elements). Each class includes one or more subclasses as the third level of the hierarchy and consists of the class symbol followed by a capital letter (e.g. H01S devices using stimulated emission). The subclass is subdivided into groups which can be main groups or subgroups. The main groups are the fourth hierarchical level in the classification while subgroups are lower hierarchical level depending on the main group. A main group symbol consists of the subclass symbol followed by a one to three digit number, the oblique stroke and the number 00 (e.g. H01S 3/00). The subgroup symbol consists of the subclass symbol followed by one to three digit number of its main group, the oblique stroke and a number of at least two digits other than 00 (e.g. H01S 3/02). The IPC is a descending hierarchical classification system; therefore the lower levels are subordinated to the higher levels. Figure 4 shows an example of a classification symbol.

Different technology classifications have been made in order to relate the IPC codes to the industrial or economic sectors. On the one hand, Institutions such as Franhoufer ISI, the *Observatoire des Sciences et des Technologies* (OST) and the French patent office (INPI) made concordance studies based on the IPC codes (Grupp & Schmoch, 1992). Later on, an improvement was prepared based on new editions of the IPC codes with support of the WIPO (Schmoch, 2008) resulting in a classification were the biotechnology field is included in the area of chemistry and pharmaceuticals, and had eight IPC codes: C07G; C12M; C12N; C12P; C12Q;

[6] Section A: human necessities; section B: performing operations, transporting; Section C: chemistry, metallurgy; section D: textiles, paper; section E: fixed constructions; section F: mechanical engineering, lighting, heating, weapons, blasting; section G: physics; section H: electricity.

C12R; C12S. Recently, in January 2013, an update made by the WIPO included the C07K to the already chosen codes[7].

**Figure 4. Complete Classification Symbol IPC, Source WIPO (2015)**




On the other hand, the OECD made a statistical framework in order to guide the measurement of biotechnology activity which includes a classification scheme (OECD, 2005). The OECD's model is based on the definition of biotechnology that involves not only modern concepts, but also traditional and borderline activities[8]. The goal of this classification was to avoid the inclusion of non-biotechnology patents and the exclusion of relevant biotechnology patents based on the analysis of the definition of biotechnology adopted by the organization. This framework has been used until our days in different documents by the OECD, and has been one of the most adopted classifications even for industries other than biotechnology (e.g. ICT) in terms of the IPC codes. Therefore, because of its specificity and broad use, the present research uses the classification of biotechnology IPC codes developed by the OECD. These IPC classes for biotechnology are (Beuzekom & Arundel,

[7] A document (Excel file) describing this fact can be consulted at the WIPO web site. The specific link is: http://www.wipo.int/ipstats/en/statistics/technology_concordance.html. This link was consulted by the author of this research in November 2013.

[8] The OECD definition of biotechnology is: "*The application of science and technology to living organisms, as well as parts, products and models thereof, to alter living or non-living materials for the production of knowledge, goods and services*." (OECD,2005)

2009): A01H1/00, A01H4/00, A61K38/00, A61K39/00, A61K48/00, C02F3/34, C07G(11/00, 13/00, 15/00), C07K(4/00, 14/00, 16/00, 17/00, 19/00), C12M, C12N, C12P, C12Q, C12S, G01N27/327, G01N33/(53*, 54*, 55*, 57*, 68, 74, 76, 78, 88, 92)[9]. A description of each of the IPC codes for biotechnology is found in Annex 1.

**6.2.1.6 Differentiating between Student and Teacher Firm in the learning dyad**

As stated in the hypothesis section, the present research differentiates between the teacher firm and the student firm in the learning dyad. Although Lane & Lubatkin initially mentioned the teacher and student firm many times in the seminal article about relative absorptive capacity (Lane & Lubatkin, 1998), there is not an explicit differentiation between both of them. The explanation of the authors is funded on the sample taken from the population of R&D alliances between pharmaceuticals and biotechnological firms. Their argument is that the emergence of biotechnology as a competence destroying technology obligated the pharmaceutical firms to create alliances with biotechnological firms in order to learn new capabilities (pp.466). Based on this, their conclusion was that the pharmaceuticals firms are students, and biotechnological firms are teachers. Nevertheless, the contrary may also occur. This is even stated by the authors when saying "*…It is important to note that the hypotheses predict one-way inter-organizational learning (i.e., one teacher and one student). This is not to say that two-way learning in alliances is rare… we believe that the factors that influence one-way learning also effect two-way learning…*" (pp.464). Consequently, the differentiation in their theoretical arguments and the call for teacher and student firms in the learning dyad doesn't appear to have any role in their results. In other words, it seems that there is not incidence in differentiating between teacher and student firms in the results.

[9] * Those IPC codes also include subgroups up to one digit (0 or 1 digit). For example, in addition to the code G01N 33/53, the codes G01N 33/531, GO1N 33/532, etc. are included..

In order to overpass the lack of operationalization with this respect, the present research intents to differentiate between both roles in the learning dyad, specially if there is a critical interest in the technological knowledge dimension under the Knowledge Based View. To do so, this study uses firms' patent portfolio as an indication of the knowledge the firm possesses. Then, following Laursen et al. (2010), the present research uses a measure of the extent of the firms' stock of knowledge grounded on the Herfindahl index applied to the patent technological capital of firms during five years before (*t-1* to *t-5*) since the date of application *t* of the patent. The index measures the amount of dispersion of the firm across three digit IPC technological classes and is defined as:

$$1 - \sum_{i=1}^{n} \alpha_i^2$$

**Equation 16  Firms' knowledge measure**

$\alpha i$ is the share of patents in three digit IPC class $i$ in the firm stock of patents. This measure ranges from 0 to 1 where 0 indicates a narrow scope of knowledge and 1 indicates a broad scope of technological knowledge and expertise. Then, comparing the indexes between both firms in the alliance result in the differentiation of the teacher or student firm depending if is a higher or lower index respectively. This research assumes that firms with a broader knowledge –higher index– may manage different technologies and therefore should have greater ability and experience that can be transferred to the student firm –lower index-.

**6.2.1.7 The collection process and data sample** Having discussed the research set up that includes industry, databases, identification tools, and firm's joint patents; it is important to address the data collection process and the data sample used in the research. There is not a direct way to obtain biotechnology joint patents or co-patents from patent databases. To do so, the first step in this process was to obtain all the biotechnology patents no matter how many owners were assigned to them. This activity was made using a standalone specialized software as an interface with the databases in order to collect the biotechnology patents using a search equation based on the IPC codes assigned to the field[10]. The resulting data from this first step had an initial cleaning of patents with incomplete information (e.g. empty fields) and patents belonging to other kind of organizations or actors different from firms[11] . Table 2 shows the number of biotechnological patents found by year before and after the initial cleaning.

The total number of firms' biotechnology patents obtained was 88797. This number of patents includes joint and single owner patents. Therefore, the second step consisted on extracting only the joint biotechnological patents from this initial database. In order to obtain these joint patents, a first text mining algorithm was design to search each of the 88797 patents and extract only those –and their information– that have more than one applicant in the applicant field. After scanning the resulting data, a second text mining cleaning algorithm was designed in order to solve two problems found: first, in some cases the same firm appeared

[10] The software used in this activity was Matheo Patent XE®. This software is designed to search, retrieve and analyze patent data from the USPTO or Espacenet databases. The used search equation was the following: ((ic:(A01H1) OR ic:(A01H4) OR ic:(A61K38) OR ic:(A61K39) OR ic:(A61K48) OR ic:(C02F3/34) OR ic:(C07G11) OR ic:(C07G13) OR ic:(C07G15) OR ic:(C07K4) OR ic:(C07K14) OR ic:(C07K16) OR ic:(C07K17) OR ic:(C07K19) OR ic:(C12M*) OR ic:(C12N*) OR ic:(C12P*) OR ic:(C12Q*) OR ic:(C12S*) OR ic:(G01N27/327) OR ic:(G01N33/53) OR ic:(G01N33/531) OR ic:(G01N33/532) OR ic:(G01N33/533) OR ic:(G01N33/534) OR ic:(G01N33/535) OR ic:(G01N33/536) OR ic:(G01N33/537) OR ic:(G01N33/538) OR ic:(G01N33/539) OR ic:(G01N33/54) OR ic:(G01N33/541) OR ic:(G01N33/542) OR ic:(G01N33/543) OR ic:(G01N33/544) OR ic:(G01N33/545) OR ic:(G01N33/546) OR ic:(G01N33/547) OR ic:(G01N33/548) OR ic:(G01N33/549) OR ic:(G01N33/55) OR ic:(G01N33/551) OR ic:(G01N33/552) OR ic:(G01N33/553) OR ic:(G01N33/554) OR ic:(G01N33/555) OR ic:(G01N33/556) OR ic:(G01N33/557) OR ic:(G01N33/558) OR ic:(G01N33/559) OR ic:(G01N33/57) OR ic:(G01N33/571) OR ic:(G01N33/572) OR ic:(G01N33/573) OR ic:(G01N33/574) OR ic:(G01N33/575) OR ic:(G01N33/576) OR ic:(G01N33/577) OR ic:(G01N33/578) OR ic:(G01N33/579) OR ic:(G01N33/68) OR ic:(G01N33/74) OR ic:(G01N33/76) OR ic:(G01N33/78) OR ic:(G01N33/88) OR ic:(G01N33/92))

[11] To take out actors different from firms from the data, the "inventor" filter was used, followed by the filter equation in the Matheo Patent XE®: not APP/univ* and not APP/empty and not APP/colle* and not APP/inst* and not APP/center and not APP/hospital*

twice in the applicant field giving the wrong idea that it's a joint patent; and secondly, in some cases the names of the inventors are different by typing errors at the inventor and applicant fields, and therefore these patents were not filtered by the specialized patent software at the initial cleaning of actors different from firms[12]. The Table 3 shows the resulted joint biotechnology patents by year after using these two text mining algorithms.

**Table 2. Raw data of biotechnology patents before and after cleaning**

| YEAR | Number of Biotech patents | Number of biotech. Patents - after cleaning |
|---|---|---|
| 2006 | 33247 | 18005 |
| 2007 | 33916 | 18947 |
| 2008 | 30987 | 17359 |
| 2009 | 29998 | 17248 |
| 2010 | 29508 | 17238 |
| TOTAL | 157656 | 88797 |

The total number of joint patents found in the biotechnological field was 1695. This set of joint patents is the population or universe that is the focus of the present research. A simple random sampling method was used to select a subset to estimate characteristics of the whole population. The resulting sample size is 465[13] joint

[12] These two algorithms were developed in MATLAB® and were designed collaboratively with Julian Mora whom was a Student of a Master in Industrial Engineering. This student was directed by the author of the current research thesis.

[13] The sample size is similar to other studies using technological distance. For example, Park, Yoon & Kim (2013) use a sample of 318 patents to support merger and acquisitions selection; Kim & Song (2007) uses 58 joint patents to understand how technology is created through alliances; Hagerdoor et al., (2003) used a sample size of 226 joint patents in their search for the effects of previous experience on joint patenting; Petruzzelli (2011) uses 796 join inventions to understand the impact of technological relatedness in the university-industry collaboration. In general,

patents using a 5% margin error, 98.8% confidence level, and a response distribution of 50%. These 465 joint patents reflect, as stated before, R&D alliance collaborations between the firms involved as applicants or owners of the technologies. Annex 2 shows the sample of joint patents and their application number.

**Table 3. Number of joint patents in the biotechnology field**

| YEAR | Number of Biotech patents | Number of joint Biotech Patents | Sample of joint Biotech Patents |
|---|---|---|---|
| 2006 | 18005 | 258 | 48 |
| 2007 | 18947 | 359 | 99 |
| 2008 | 17359 | 339 | 107 |
| 2009 | 17248 | 372 | 110 |
| 2010 | 17238 | 367 | 101 |
| TOTAL | 88797 | 1695 | 465 |

The 465 joint patents are owned by 284 firms. Information from these joint patents and firms was extracted in order to set the different variables in the research. For instance, this sample of joint patents was cited (forward citation) by other 777 patents, and cited (backward citation) by 7505 patents. Furthermore, the use of information from each of the firms owning the joint patents was analyzed to obtain the firm's technological profiles in order to measure the relative technological distances and their technological capital. It is important to highlight that information about each firm was searched in the databases of *Bloomberg®*, *Lexis Nexis®,* and *U.S Edgar Securities and Exchange Commission* to know about relation to subsidiaries, merges, acquisitions and name changes to avoid apparent alliances

differences are dependent of elements such as the access to different types of specialized and sometimes paid databases, the matching between more than one databased used in the design of the study, the cost of processing or transforming the data, or the use of specialized tools that include information already transformed or processed (e.g. backward or forward citations).

from the same firm. This information also has an important effect on the different variables of the research. The next section explains with more detail the variables definition and operationalization.

**6.2.2 Variable Definitions And Operationalization** In this section, the variables of the research are discussed. Firstly, the dependent variable of the study is discussed. Next, the independent and control variables are introduced. The independent variables of the study are collected for the years 2001 to 2010. The dependent variables are collected from 2006 to 2014. This is described in detail in the following sub-sections. An overview of the main variables with their descriptions is provided in Table 4.

**6.2.2.1 Dependent Variable**

- **Value of joint invention:** Patents are good indicators of innovation performance overpassing some other indicators such as R&D spending (Griliches, 1990). Likewise, joint patents have been used in research studies as the main innovation output of collaborative R&D in alliances (C. Lin et al., 2012; Sampson, 2007). This research uses joint patent applications as the innovation outcome in collaborative R&D alliances and measure their value as patent counts to derive the dependent variable of the study (*JointValue*). Value estimates of patents is difficult to make because as known since early studies their market is not totally open and manages asymmetric information increasing the uncertainty in the estimation analysis (Schankerman & Pakes, 1986).

  In general, there is a distinction between the private value and the social value of inventions or patents (Baron & Delcamp, 2012). On one hand, the private value could be defined as the benefits perceived by the firm that has the patent (e.g. difference in profits) compared to the firm that does not have it (Harhoff, Scherer, & Vopel, 2003). In general, to estimate the private value of a particular patent the following methods are used which take into account the value added of its owner: based on costs, market, or discounted cash flow (WIPO, 2005).

**Table 4. Definition of dependent, independent, and control variables**

| | |
|---|---|
| **Dependent Variable** | |
| JointValue | Number of citations received by each joint patent application in the next five years from the application date. |
| **Independent Variables** | |
| TeachDist | The complement of the Jaffe cosine similarity measure between the teacher firm technological profile and the profile of the joint invention in the alliance. |
| StuDist | The complement of the Jaffe cosine similarity measure between the student firm technological profile and the profile of the joint invention in the alliance. |
| TeachCap | Count of the number of patents that a teacher firm successfully filed during the previous five years to the application date. |
| StuCap | Count of the number of patents that a student firm successfully filed during the previous five years to the application date. |
| **Control Variables** | |
| GeoDist | Natural logarithm of the physical distance expressed in kilometers between the location sites of firms jointly developing a patent. |
| PriorColl | Number of registered joint patents between the firms developing a joint patent in the alliance in the previous five years to the application date. |
| NumCod | Number of International Patent Classification codes (IPC) assigned to the joint patent under analysis. |
| PatCit | Number of patent backward citations referenced in the joint patent under analysis. |
| PCT | Dummy variable set to one if the joint patent under analysis is using the Patent Cooperation Treaty. |
| Years | Dummy variable indicating a particular year in the observed period 2005-2010 |

The costs method uses past information about the costs to create, develop, protect and even commercialize the technology; the method based on the market made a comparison with a similar technology in the market in order to approximate de value; and the method based on the discounted cash flow is grounded on the Net Present Value of the technology.

On the other hand, the social value embodies the total net value generated by the patent for social welfare (Baron & Delcamp, 2012). The social value of a patent commonly uses indicators (Harhoff, Scherer, & Vopel, 2004; Reitzig, 2003, 2004) that, as evidence has shown, are correlated to their private value (Y. G. Lee, 2009; Tseng, Hsieh, Peng, & Chu, 2011). The most used indicators in the patent value analysis are: number of forward citations, number of family members of a patent (Gambardella, Harhoff, & Verspagen, 2008), and breath of a patent (Harhoff et al., 2003). The forward citation method is based on the supposition that if a patent is the building block of further inventions, then the value of the exclusion right increases. The family size assumption captures the number of jurisdictions in which a patent is protected as a sign of increase on value taking into account the territoriality of patents. Finally, the breath or scope of a patent assumes that patents that are included in many product or processes increase the value of the right of exclusion.

In general, forward citation is a reflection of the quality of the patent and has been largely adopted by researchers when analyzing indicators to examine patent value (Acosta, Coronado, & Martinez, 2012; Makri et al., 2010; Petruzzelli, 2011; Phene et al., 2006; Schettino, Sterlacchini, & Venturini, 2013; Singh, 2008). This indicator has been used already valuing joint patents (Briggs & Wade, 2014; Briggs, 2015; Petruzzelli, 2011) and has shown more significance over other indicators of patent value (Fischer & Leidinger, 2014). The present study measures the value of R&D alliance outcomes as the forward citation of the joint patent within five years of the application date, excluding self-citation of the firm owners of the joint patent.

Forward citation is defined as the number of cites the joint patent receives in subsequent patents and reflects the ability of the joint patent to support future inventions and the incentive of continual patents (Makri et al., 2010). To measure the number of cites received by a patent, it has been recommended to use short citation spans (Lanjouw & Schankerman, 1999). This might be considering that organizational memory depreciates promptly (Argote, 2013), and the knowledge capital depreciates most of its value within approximately five years (Griliches, 1979). Although some studies have used time windows of three years (Briggs & Wade, 2014; Briggs, 2015)[14], or four years (Sampson, 2007) this research uses a time window of five years (Mariani & Romanelli, 2007; Schettino et al., 2013) to control the bias of early patents having more citations since they are being cited for longer periods. Then, the variable (*JointValue*) is measured by the number of citations from the year $t$+1 to the year $t$+5 that a joint patent application of year $t$ receives from subsequent patents.

**6.2.2.2 Independent Variables** The independent variables used in the research are the student firm technological distance (*StuDist*), the teacher firm technological distance (*TeachDist*), and the technological capital of both, the student (*StuCap*) and teacher firms (*TeachCap*).

- **Relative technological distances:** As explained in the hypotheses section the relative technological distance is the distance from the technological knowledge embedded in the joint invention to the technological knowledge base of any of the owner firms of the patent or '*learning dyad*' (Lane & Lubatkin, 1998). Then, there are two relative technological distances in the learning dyad associated to the two firms involved in the joint invention: the student firm technological distance, and the teacher firm technological distance.

[14] Briggs & Wade (2013) tests the sensitivity of truncation using the citation of the patent over the life of the patent finding similar overall results to those using three years.

In order to measure the relative technological distances, this research used Jaffe's measure of similarity (Jaffe, 1986). This measure of technological distance makes use of vectorized patent data and might be one of the most used measures in the technological distance literature. Thus, the relative technological distance in this research is defined as:

$$Relative\ Tech.Distance_{ij} = 1 - \left( \frac{f_i f_j'}{\sqrt{(f_i f_i')(f_j f_j')}} \right) = 1 - \left( \frac{\sum_{k=1}^{n} f_{i,k} \cdot f_{j,k}}{\sqrt{\sum_{k=1}^{n} f_{i,k}^2} \sqrt{\sum_{k=1}^{n} f_{j,k}^2}} \right)$$

**Equation 17 Jaffe Cosine of Similarity measure**

In this measure, $fi$ is a vector formed only by the joint patent $i$ and $fj$ is a vector formed by all the patents or patent portfolio registered by its owner firm $j$; both of them $fi\ and\ fj$ are allotted to the $k$ patent class. The term inside the parenthesis is the dot product or cosine of the angle between both vectors and is also known as cosine similarity measure which is an un-centered Pearson correlation between the two vectors. Then, if the two vectors $fi\ and\ fj$ coincide perfectly the parenthesis term takes the value one; but if they do not overlap at all, the vectors are orthogonal and the value is 0.

To allocate the technological classes $k$ to each vector $fi\ or\ fj$ the present study uses the classes at the three digit level (Nambisan, 2013; Nooteboom et al., 2007) assigned to each patent by the technological codes of the International Patent Classification –IPC- administered by World Intellectual Property Organization –WIPO. In contrast to the use of specific classes from the OECD as was done to obtain the biotechnology industry research sample, here all the classes are considered without exception in order to obtain a technological profile from the technological knowledge bases embedded in the joint patent $fi$

and the owner firms $fj$ of the learning dyad. Therefore, vectors $fi$ $and$ $fj$ are located in technological space with a dimension represented by the 129 patent classes derived from the three digit level of the IPC codes.

The technological profiles of $fi$ $and$ $fj$ are built by counting the number of patents in each patent class $k$ using the IPC coding. Then, on the one hand the joint patent profile is simple to calculate because it can have at most one IPC code per class at the year of the application. On the other hand, the owner firms' technological profile is more complex to calculate because it can have as many IPC codes per class as many patents they have. Then, a time window is necessary to control the number of patents applied by the firms (e.g. newer firms could have less patents than older firms). Therefore, to calculate the technological profile of each one of the 284 firms it was necessary to obtain the whole patents from each firm during a time moving window of five years from $t-1$ to $t-5$ being $t$ the year of application of the joint patent. This has been done on past research to calculate technology profiles (Nooteboom et al., 2007; Petruzzelli, 2011; van de Vrande et al., 2011). Although some studies use three year time windows (Schildt et al., 2012), the five year time was selected taking into account that knowledge capital depreciates rapidly (Griliches, 1979) and that technological impact is best assessed during this time (Ahuja, 2000; Henderson & Cockburn, 1996).

In the present research a third text mining algorithm[15] was developed in order to create the technological profiles. First, it was necessary to collect and prepare the input data from the databases which consisted of both, the technological IPC classes assigned to each of the joint patents, and the technological IPC classes assigned to each one of the patents belonging to the owner firms since

[15] This third algorithm was developed in MATLAB® and was designed collaboratively with Julian Mora as a Student of a Master in Industrial Engineering. This student was directed by the author of the current research thesis.

five years before the application date of the joint patents. Second, the algorithm calculates the technological profile of joint patents classifying and signaling technology classes to obtain a vector with the classes the joint patent includes (Figure 5). Besides, the technological profile of owner firms is calculated classifying and signaling the technology classes of each patent belonging to the portfolio of owner firms, and then adding up all the patents to obtain a final vector that includes all the classes assigned to the patents (Figure 6).

**Figure 5. Example of technological Profile of the joint patent**

| Technology Class | Joint Patent Technology Profile |
|---|---|
| *1 (e.g. A01H1/00)* | 1 |
| *2 (e.g. C12Q1/70)* | 1 |
| *3 (e.g. C07K1/04)* | 0 |
| . | . |
| . | . |
| . | . |
| . | . |
| *k* | 1 |

**Figure 6. Example of a technological profile of an owner firm**

| | Firm Technology Patents | | | | | | | | | Firm Technology Profile |
|---|---|---|---|---|---|---|---|---|---|---|
| Technology Class | Patent 1 | | Patent 2 | | Patent 3 | . . . | | Patent n | | |
| *1 (e.g. A01H1/00)* | 1 | + | 0 | + | 1 | . . . | + | 1 | = | 7 |
| *2 (e.g. C12Q1/70)* | 1 | + | 1 | + | 1 | . . . | + | 0 | = | 12 |
| *3 (e.g. C07K1/04)* | 0 | + | 0 | + | 0 | . . . | + | 1 | = | 21 |
| . | . | + | . | + | . | . . . | + | . | = | . |
| . | . | + | . | + | . | . . . | + | . | = | . |
| . | . | + | . | + | . | . . . | + | . | = | . |
| *k* | 0 | + | 1 | + | 1 | . . . | + | 0 | = | 3 |

Third, once the technology profiles of owner firms and joint patents were obtained, it was necessary to assembly which owner firms were attached to which joint patents. To do so, an adjacency matrix was created. An adjacency matrix is a matrix in which the rows and columns represent nodes and an entry in row $i$ and column $j$

represent a tie or linked from $i$ to $j$ (S. P. Borgatti, Everett, & Jhonson, 2013). In this research the matrix is symmetrical and the nodes are represented by the joint patents or firm owners. Then, the adjacency matrix A in which $aij$=1 means that there is a tie or link from $i$ to $j$, otherwise $aij$=0. However, because the matrix is symmetrical, then $aij$=$aji$, where row $i$ is a joint patent and column $j$ is an owner firm, or vice versa. Figure 7 shows an example of an adjacency matrix.

**Figure 7. Example of adjacency matrix**

| | JP 1 | JP 2 | JP 3 | . | . | . | JP m | Firm 1 | Firm 2 | Firm 3 | | | Firm n |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| JP 1 | 0 | 0 | 0 | . | . | . | 0 | 1 | 0 | 0 | . | . | 1 |
| JP 2 | 0 | 0 | 0 | . | . | | 0 | 0 | 1 | 1 | . | . | 0 |
| JP 3 | 0 | 0 | 0 | . | . | . | 0 | 0 | 1 | 0 | . | . | 0 |
| . | . | . | . | | | | . | . | . | . | | | . |
| . | . | . | . | | | | . | . | . | . | | | . |
| . | . | . | . | | | | . | . | . | . | | | . |
| JP m | 0 | 0 | 0 | . | . | . | 0 | 1 | 0 | 1 | . | . | 1 |
| Firm 1 | 1 | 0 | 0 | . | . | . | 1 | 0 | 0 | 0 | . | . | 0 |
| Firm 2 | 0 | 1 | 1 | . | . | . | 0 | 0 | 0 | 0 | . | . | 0 |
| Firm 3 | 0 | 1 | 0 | . | . | . | 1 | 0 | 0 | 0 | . | . | 0 |
| . | . | . | . | | | | . | . | . | . | | | |
| . | . | . | . | | | | . | . | . | . | | | |
| Firm n | 1 | 0 | 0 | . | . | . | 1 | 0 | 0 | 0 | . | . | 0 |

In the example of Figure 7 joint patents (JP) and owner firms (Firm) are ordered in the rows and columns of the matrix. A value of 1 in any cell means that the row and column of that cell are linked. For example, in Figure 7 Firm 1 is linked to joint patent 1, and Firm 2 and Firm 3 are linked to join patent 2[16]. As was

[16] In this case of the example, Firm 2 and Firm 3 are the "dyadic learning" involved in the relative technological distance to the joint patent 2.

stated, the matrix is symmetrical to the main diagonal because the relation is not directed (e.g. a join patent is linked to a firm or the firm is linked to the joint patent). The adjacency matrix was obtained using UCINET 6.0 software (S P. Borgatti, Everett, & Freeman, 1999). Ucinet is a matrix oriented software package for the analysis of social network data with routines that facilitates the creation of adjacency matrices.

Finally, having the technological profiles of joint patents and owner firms, and the adjacency matrix that shows the links between them, a fourth algorithm was created to calculate the relative technological distance following the Jaffe measure described above. The algorithm identifies in the adjacency matrix, the links between the joint patent and owner firms, then looks for the technological profiles stored vectors of both of them and calculates their relative technological distance. Then the two relative technological distances of the two owner firms associated to the single joint patent are classified as a student (*StuDist*) or the teacher (*TeachDist*) firm according to the breadth of knowledge of the firms.

- **Technological capital**: As stated at the hypotheses section, intellectual capital is the ability to convert new ideas into products or services (Booth, 1998), or in other words, to convert invisible assets such as knowledge into resources (Bradley, 1997). Then, intellectual capital can be seen as knowledge transformed into technology (Hsu & Wang, 2012). Then, technology is an output of the intellectual capital of firms which can be represented by patents. Consequently, patents are indirect measures of the technological capital of firms (Narin, Noma, & Perry, 1987). In this research, technological capital of the student (*StuCap*) and teacher (*TeachCap*) firms are measured by the cumulative patents owned by them.

  The technological capital is a representation of the R&D capabilities of firms (Nooteboom et al., 2007). Patents granted to a company are used to measure the technological capabilities of firms (Narin et al., 1987) and therefore their technological capital (Petruzzelli, 2011). Specifically in this study, cumulative technological capital of student and teacher firms (*StuCap* and *TeachCap* respectively) is calculated as the number of patents that a firm obtained in the 5 years prior (*t-1* to *t-5*) to the issue date of the joint patent application *t* (Stuart, Ozdemir, & Ding, 2007; van de Vrande et al., 2011; Veugelers & Cassiman, 2005).

**6.2.2.3 Control Variables** The present research includes some variables to control for the alternative factors that can explicate the value of innovations jointly developed by firms. On the one hand, firm's relational control variables are introduced: geographical distance and prior collaboration. On the other hand, variables that can affect directly the value of invention are presented: backward citation, PCT, technological scope, and years.

- **Geographical Distance**. Physical distance between firms has an effect in the creation of knowledge (Greunz, 2003) and therefore are determinants of R&D collaboration and innovation (Cunningham & Werker, 2012; Scherngell & Barber, 2009). Geographical distance can be a reinforcing dimension to its technological counterpart (Mattes, 2012) and has shown a relationship to innovation value (Petruzzelli, 2011). The current research uses the distance in kilometers between firms' countries in order to control for the fact that the partners of the joint invention originate in a closer or farther proximity. Therefore, in this study the variable geographical distance (*GeoDist*) is measured as the natural logarithm of the physical distance in kilometers between countries were the firms belongs to. However, a modification in the distance of firms located in the United States was made noticing that the distance between States inside this country could be larger than distances between pair of countries (e.g. European Countries). For instance, the distance between the states of California and Massachusetts (4112.54 Km) is higher than the distance between the countries France and Germany (815.79 Km). And because the patent database used is the United States office, many of the joint patents are conformed from firms from this same country. Consequently, geographical distance between states in United States was measured in order to avoid a possible bias.

Firms choose their partners based on how close they are (Broekel & Boschma, 2012). On the one side, closeness between firms have low technology transfer costs and is more effective in coordinating joint technology (C. Kim & Song, 2007). Besides, physical proximity helps to the face-to-face interactions (Antonelli, 2000),

cooperation, trust, and tacit knowledge transfer (McKelvey, Alm, & Riccaboni, 2003) favoring the creation of common practices important to innovation and knowledge spillovers (Phene et al., 2006). On the other side, because of the difficulty presented on the knowledge flows between actors at high physical distances, the geographical distance can have a negative effect on the probability of collaboration of firms (Scherngell & Hu, 2011). This can bring the idea that knowledge flows are geographical localized (Jaffe, Trajtenberg, & Henderson, 1993) where learning is better if distance is smaller (Verdolini & Galeotti, 2011). Based on what have been explained, it is expected that firms within more proximity have a positive effect in the dependent variable *JointValue*. In other words, high geographical distances have a negative effect in *JointValue*.

- **Prior Collaboration**. Prior collaboration between firms influence the choice of future partners (C. Kim & Song, 2007), lowers transactional costs (Dyer & Chu, 2003), and creates trust (Hagedoorn et al., 2003; Petruzzelli, 2011). The prior collaboration and the ongoing interaction make partners learn from each other needs and capabilities (Gulati, 1995a), and therefore improve the prediction of collaboration patterns and behavior (Dyer, 2002). In addition, collaborative behavioral ambiguity might be reduced because the information asymmetries in the alliance partners decrease over time (Casciaro, 2003). Consequently, knowledge sharing routines increase over time (Dyer & Singh, 1998; Inkpen & Dinur, 1998) leading to simplify communication and joint problem solving (Yli-Renko, Autio, & Sapienza, 2001).

Knowledge sharing routines are structures embedded in the knowledge transfer process across the companies (Dyer & Hatch, 2006; Dyer & Singh, 1998). Then, learning between partners makes these structures also improve over time, and allows a better absorption and exploitation of external knowledge from partners (Simonin, 1999). Therefore, these relational routines facilitate the sharing (Kogut, 1988), joint development (Petruzzelli, 2011), and coordination of partner's particular technologies (Kale & Singh, 2000). As a result, it is estimated that theV variable prior ties (*PriorColl)* has a positive effect on the dependent variable *JointValue*. To measure prior ties between the firms in the "learning dyad", this research checked for past joint patents between the partner firms. Going further than Petruzelly (2011), this research uses the variable *(PriorColl)* counting the number of joint patents the same partners own in the five years prior to the issue date of the joint patent application under analysis.

- **Backward Citations.** This variable represents the past activity exploration of the firm and therefore the ability to monitor the technological surrounding knowledge (Laursen et al., 2010). In other words, backward citations measures the technological search made by the firm in order to build on it the new knowledge (Katila & Ahuja, 2002). The number of cited references in patents have been already used in some studies resulting in a positive effect in innovation value (Gambardella et al., 2008; Harhoff et al., 2003; Sneed & Johnson, 2009). It is expected that the larger the number of references, the larger the technological knowledge input, and thus, the larger the probability to lead to more valuable innovative outputs (Gay & Le Bas, 2005; X. Hu, Rosseau, & Chen, 2012). For that reason, the present research estimates that the number of patents cited as prior art or patent citations –*PatCit*– has a positive effect on the joint invention value of the alliance –*JointValue*–.

- **Patent Cooperation Treaty.** The Patent Cooperation Treaty (PCT) is a strategy to achieve international patent protection. Many research studies have confirmed that the decision to use PCT for their patent application is an indicator of patent value (Guellec & Van Pottelsberghe de la Potterie, 2002; Reitzig, 2004). Accordingly, this research assumes that if a patent is using the PCT strategy in order to protect its intellectual property rights in other countries, then the firm owning that patent is confident that the technology embedded in the patent will represent an economic or strategic return that has to be even higher than the cost invested in the process itself. Therefore, the patent application should be considered as having high value. The present research uses a dummy variable called *PCT* in order to establish if the joint patent has applied for the Patent Cooperation Treaty (*PCT=1*) or not (*PCT=0*). It is expected that if variable *PCT=1,* then the value is higher than if variable *PCT=0*.

- **Technological Scope**. The scope of a patents is an important element of the efficacy of patent protection (Scotchmer, 1991). From past research, it’s known that the scope of technological classes has a positive effect on the quality of patents (Lanjouw & Schankerman, 1999). As a consequence, technological scope has been used as a correlated variable with value in order to estimate the patent value and its determinants (Harhoff et al., 2003). The International Patent Classification codes (IPCs) capture the number of knowledge areas in which a patent for a single invention is related to. Following Lerner (1994), the present research relates the value measure to the number of IPCs -*NumCod*- cited in the patent, assuming that this variable captures the scope of the patented invention. This research measure the scope of the joint patented technology as the number of different IPC classification codes in the application document. Hence, it is expected that the number of IPC codes, *NumCod*, in the application document has a positive effect on the joint invention value of the alliance *JointValue*.

- **Year variable**. Year dummy variables are used as control variables in order to reflect potential exogenous effects –changes over time– characterizing the year of the joint patent application that probably influence the propensity of patents to be cited. In general, yearly dummy variables has been used in studies related to innovation performance (Gilsing et al., 2008; Nooteboom et al., 2007) and innovation value (Petruzzelli, 2011).

**6.2.3 Model specification and estimation** The hypotheses developed in Chapter 3 identify a dependent variable, value of joint invention (*JointValue*). Hypotheses 1 and 2 predict how the relative technological distances in the learning dyad –student and teacher– affect the dependent variable.

Hypothesis 3 predicts how technological capital of firms in the learning dyad –student and teacher– affects the dependent variable. Empirical validation of the hypothesis is done through the examination of 465 joint patent applications over the period of 2006 to 2010. In general, the model explaining joint invention value can be written as the following equation:

$$JointValue = fn\ (StuDist;\ TeachDist;\ StuCap;\ Teach\ Cap;\ control\ variables)$$

**Equation 18 Model specification of the variables**

The main empirical method applied in this research is the regression analysis. The dependent variable is regressed against a vector of explanatory variables which includes both hypothesized explanatory and control variables. The present study used a first approach using a Poisson distribution. However, based on the limitation of overdispersion of the Poisson model, the Negative Binomial regression is better suited for modeling this count data as described in the following sections.

**6.2.3.1 First Approach: Poisson Model** In statistics, count data refer to observations with only non-negative integer values from zero to some undetermined value, and a count variable is a specific list of count data (Hilbe, 2014). The dependent or response variable of this study is a non-negative and integer count variable, and therefore violates the assumptions of homoscedastic and normally distributed errors. The Poisson regression is a first approach that offers an expected baseline model for such data (Ahuja, 2000; Hausman, Hall, & Griliches, 1984; Henderson & Cockburn, 1996). The Poisson distribution has a criterion of equi-dispersion because the single parameter to be estimated, $\mu$ or the mean, is the same variance as shown in the following equation.

$$f(y;\mu) = \frac{e^{-\mu_i}(\mu_i)^{y_i}}{y_i!},\ y_i = 0,1,2,\ldots n_i;\ \mu > 0$$

**Equation 19 Poisson distribution**

In order to obtain the regression model from the Poisson distribution, it is derived by letting $\mu$ to depend on the regressors and parameters. Therefore, $yi$ given $\boldsymbol{X}i$ is Poisson distributed with density:

$$f(y_i/\boldsymbol{X}_i) = \frac{e^{-\mu_i}(\mu_i)^{y_i}}{y_i!},\ y_i = 0,1,2,\ldots n_i;\ \mu > 0$$

**Equation 20 Density of Poisson distribution**

Here, $\boldsymbol{X}i$ is a k-dimensional vector of regressors. If $\boldsymbol{\beta}$ are the parameters, the mean in the model is parameterized as:

$$\mu_i = e^{(X_i'\beta)} \text{ or } ln\, \mu_i = \boldsymbol{\beta X}_i$$

**Equation 21 Poisson regression model**

As a result, the Equations 20 and 21 define the Poisson regression model for the count data (Cameron & Trivedi, 2013). However, most of real count data models suffered of overdispersion: variability of data is greater than the mean. This criterion plays an important role in the modeling of count data for the reason that if the model is overdisperse the standard errors are biased and cannot be trust, then a variable may give the impression to be a significant predictor when it's not. The count data of the present research seems to be overdispersed as shown in Equation 22:

$\mu$=1.670968 ; $var$=12.022970

**Equation 22 Mean and variance of variable *JointValue***

Nevertheless, in a more general sense, overdispersion occurs when the observed variance of the Poisson count response is greater than the variance of the predicted or expected counts (Hilbe, 2014). Then, it is critical to test for overdispersion the count data models based on Poisson distributions (Cameron & Trivedi, 1990). After the model was tested (Appendix 3, Section 3.1[17], Section 3.2), the following results were obtained:

Pearson $x2$= 1927.497 ; $DS$=4.2738 ; Zero Counts = 57.6344086 %

**Equation 23 Indicators of Poisson overdispersion**

[17] The statistical analysis of the present research was realized using the software R®. The scripts as R Markdown Language are found at the Appendix 3 at the end of the document.

The Poisson model has a dispersion parameter (DS) higher than one (DS>>1), excessive zero counts in the dependent variable *JointValue* given its mean 1.67[18], and an observed variance higher than its mean. The observed variance is 12.0229, while the expected variance is 1.670968 indicating a considerable overdispersion (Appendix 3, Section 3.3). A considerable difference is found between the Standard Errors and the Scaled Standard Errors by the DS parameter (Appendix 3, Section 3.4). Even further, the values of Robust Standard Errors differ substantially from Model Standard Errors, showing extra evidence that the count model is extradisperse (Appendix 3, Section 3.4). In conclusion, there is strong evidence to argue that the Poisson Model is overdisperse. Hence, because is not a well-fitted Poisson model, notwithstanding of the significance of the predictors (Appendix 3, section 3.4), a new model needs to be proposed.

**6.2.3.2 Negative Binomial Model** An alternative for adjusting Poisson overdispersion is the Negative Binomial model (Hausman et al., 1984). The negative binomial model does not assume the mean-variance equality of the count-dependent variable, and accounts for omitted variable bias while concurrently estimating heterogeneity. This model has been already used in studies related to innovation performance and value that presented Poisson overdispersion (C. Kim & Song, 2007; Petruzzelli, 2011). The negative binomial is derived as a Poisson-Gamma mixture model with a dispersion parameter being distributed as gamma shaped (Hilbe, 2014). The negative binomial probability distribution is expressed as:

$$f(y;\mu,\alpha) = \frac{\Gamma(y_i + \alpha^{-1})}{\Gamma(\alpha^{-1})\Gamma(y_i + 1)}\left(\frac{1}{1+\alpha\mu_i}\right)^{\alpha^{-1}} \left(1 - \frac{1}{1+\alpha\mu_i}\right)^{y_i}$$

**Equation 24 Negative Binomial probability distribution**

[18] Given a mean of 1.67, it is expected that 18% of the observations in the Poisson model have a zero count. (Appendix 3, Section 3.2)

In the above equation, $\Gamma$ is the Gamma function and $\alpha$ is the dispersion parameter. Restructuring the Gamma functions to the form of a combination results in another popular expression of the negative binomial probability distribution (Hilbe, 2011):

$$f(y;\mu,\alpha) = \binom{y_i + \frac{1}{\alpha} - 1}{\frac{1}{\alpha} - 1} \left(\frac{1}{1+\alpha\mu_i}\right)^{\alpha^{-1}} \left(\frac{\alpha\mu_i}{1+\alpha\mu_i}\right)^{y_i}$$

**Equation 25 Combination form of the Negative Binomial probability distribution**

The dispersion parameter $\alpha$ is a measure of the modification required to accommodate the additional variability, or heterogeneity, in the data. Therefore, values of $\alpha$ greater than zero specify that the model has adjusted for overdispersion and when it is zero (0) the model is Poisson[19]. The results from the negative binomial model are shown in the next section.

[19] In the Poisson model the mean and variance are equal to μ; while in the Negative Binomial model the $Mean$=μ and the $Variance$= μ(1+ $\alpha$μ)=μ+$\alpha$μ$^2$

# 7 RESULTS

This section contains the results from testing the hypotheses. It presents the main results of the dissertation –the findings of the effect of the relative technological distances of the firms in the learning dyads on the invention value of the alliance, and the findings of the effects of the technological capital of the firms in the learning dyads on the invention value of the alliance. In the first section, the descriptive statistics is presented, while in the second section the results from the independent variables and control variables in the Negative Binomial Model are presented. The key results of the dissertation are summarized in the last section.

## 7.1 DESCRIPTIVE STATISTICS

The sample dataset comprises 465 joint patents engaging 284 firms over the period 2006 to 2010. The sample represents 20 countries[20]. From the sample, 315 joint patents are collaborative efforts between firms from the same country. As Figure 8 shows, the majority of joint patents (23.5%) occurred in the year 2009, and the fewest number of joint patents was found in 2006 (10.3%). A significant rise of joint patent activity from 2006 to 2007 and a subsequent stabilization from 2007 to 2010 are clearly visible.

[20] United States, Japan, Great Britain, Switzerland, Germany, Denmark, France, Belgium, Israel, Netherlands, Canada, Australia, Italy, Bermuda, Korea, Argentina, Finland, Singapore.

**Figure 8. Sample distribution by year**

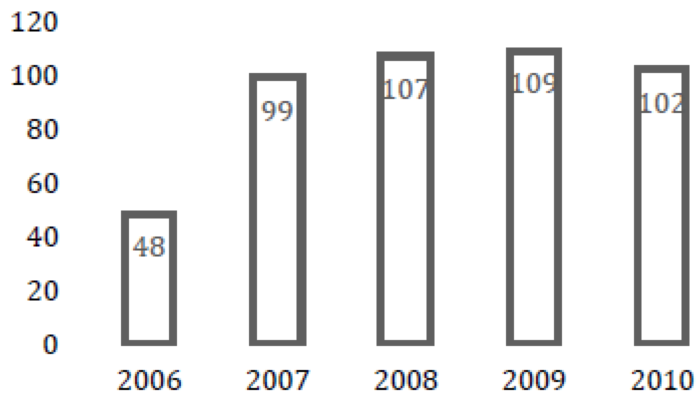



Table 5 contains the correlation and descriptive statistics information, including the minimum values, maximum values, means, and standard deviations of the variables (Apendix 3, Section 2.3). The average cites received by the joint inventions is 1.67. Figure 9 shows the forward citation distribution of the joint patents. Most of the joint patents from the sample (57.3%) are not cited at all, 13% is cited once, 7.5% is cited twice, and 5.3% is cited three times. 3.4% of the joint patents is cited more than 10 times, 1.3% is cited more than 16 times, and only 0.6% is cited more than 20 times. This Figure shows how skewed is the distribution of the count dependent variable.

**Figure 9. JointValue frequency distribution**

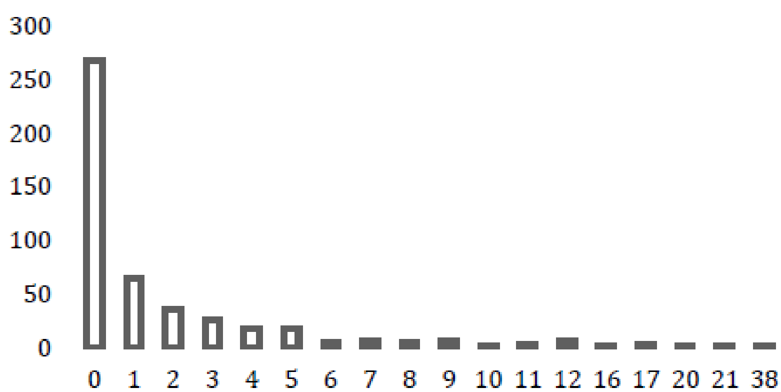


**Table 5. Correlation and descriptive statistics**

| | JointValue | TeachDist | StuDist | TeachCap | StuCap | GeoDist | PriorColl | NumCod | PatCit | PCT |
|---|---|---|---|---|---|---|---|---|---|---|
| JointValue | 1 | -0.035 | -0.073 | -0.195 | 0.140 | -0.005 | -0.243 | -0.021 | 0.124 | 0.204 |
| TeachDist | -0.035 | 1 | 0.404 | -0.031 | -0.141 | -0.059 | -0.328 | -0.222 | -0.027 | 0.452 |
| StuDist | -0.073 | 0.404 | 1 | -0.093 | -0.075 | -0.001 | -0.191 | -0.187 | -0.004 | 0.240 |
| TeachCap | -0.195 | -0.031 | -0.093 | 1 | -0.088 | 0.082 | 0.338 | -0.069 | -0.065 | -0.122 |
| StuCap | 0.140 | -0.141 | -0.075 | -0.088 | 1 | 0.014 | -0.024 | 0.140 | 0.086 | -0.001 |
| GeoDist | -0.005 | -0.059 | -0.001 | 0.082 | 0.014 | 1 | 0.087 | 0.061 | -0.073 | -0.037 |
| PriorColl | -0.243 | -0.328 | -0.191 | 0.338 | -0.024 | 0.087 | 1 | 0.086 | -0.214 | -0.499 |
| NumCod | -0.021 | -0.222 | -0.187 | -0.069 | 0.140 | 0.061 | 0.086 | 1 | 0.109 | -0.337 |
| PatCit | 0.124 | -0.027 | -0.004 | -0.065 | 0.086 | -0.073 | -0.214 | 0.109 | 1 | 0.070 |
| PCT | 0.204 | 0.452 | 0.240 | -0.122 | -0.001 | -0.037 | -0.499 | -0.337 | 0.070 | 1 |
| N | 465 | 465 | 465 | 465 | 465 | 465 | 465 | 465 | 465 | 465 |
| Mean | 1.671 | 0.367 | 0.316 | 60.877 | 38.267 | 6.168 | 26.966 | 5.082 | 16.140 | 0.471 |
| St.Dev | 3.467 | 0.298 | 0.312 | 45.405 | 34.006 | 3.091 | 44.831 | 5.266 | 30.004 | 0.500 |
| Min | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| Max | 38 | 1 | 1 | 229 | 139 | 9.629 | 267 | 39 | 225 | 1 |

On the one hand, as expected, the average of the teacher firm technological distances is higher than the student firm technological distances (0.37 and 0.32 for the teacher and student firm respectively). This can be explained taking into account that the broader the knowledge firms have accumulated, the more technological distance they can manage from their technological portfolio (Laursen et al., 2010). However, the average of the relative technological distances show, in general, that the learning dyad is cautious in terms of going into high distant technological innovation and that the difference between both of them is not large. On the other hand, on average, the teacher firm has a higher technological capital than the student firm. This result was also expected because of the direct relation between the absorptive capacity and the accumulated knowledge base that the firms possess (Cohen & Levinthal, 1990): the more experienced the firms, the more accumulated knowledge is expected to be found.

At first glance of the correlation information, innovation value is negatively impacted by the relative technological distances: increasing the relative technological distances generates a decrease on the innovation value. Besides, the correlation table shows a medium-high correlation index between some of the variables (e.g. *TeachDist-StuDist*, *TeachDist-PCT*, *PCT-PriorColl*) that can potentially lead to multicollinearity in the model. Although some studies relied their analyses on coefficients bellow a recommended *r=0.8* (King, Slotegraaf, & Kesner, 2008; Stellner, 2015) or *r=0.7* threshold, specifying acceptable discriminant validity (J. Cohen, Cohen, West, & Aiken, 2003; Petruzzelli, 2011) and the correlation coefficients between variables in Table 5 are bellow this thresholds; the correlation information is not a definitive indicator in order to decide about multicollinearity in the model. Therefore, formal tests were made to improve the understanding about this matter. Tests involving the variance inflation factors (VIFs) and condition index (CI) were made to check potential multicollinearity among the research variables. According to the tests (Appendix 3, Section 2.5), the higher individual VIF score is 2.671201, and the mean VIF score is 1.7161. Besides, the regression collinearity diagnostic procedure of Belsley (Belsley, Kuh, & Welsch, 2004), was made to examine the matrix of independent variables and obtain the condition indexes –CI–. The maximum condition index obtained had a value of 14.438. Because the variables do not have VIF values higher than 10 (J. Cohen et al., 2003) and CI values higher than 30 (Belsley, 1991) the multicollinearity of variables seems not to be a problem in the variables and data analysis in the model.

## 7.2 NEGATIVE BINOMIAL MODEL

The results of the negative binomial regression are reported in Table 6. Overall, the models support a good fit of the negative binomial regression with our data. Model 1 represented in the first column includes only the control variables (*GeoDist, PriorColl, PatCit, PCT, and NumCod*) as a base case to compare the results against, Model 2 introduces the Teacher Relative Technological Distance (*TeachDist*) to investigate its effect on joint innovation value (*JointValue*). Model 3 introduces the Student Relative Technological Distance (*StuDist*) variable. Model 4 and Model 5 include the Teacher and Student Technological Capitals (*TeachCap* and *StuCap* respectively) from the learning dyad. Model 5 take into account the simultaneous effects exerted by the variables under analysis. Model 6 also include the effects of all the variables but using robust variance estimators. There is no appreciable difference of Model 6 with the standard errors of Model 5. For the sake of simplicity the discussion is based on the Model 6, the full model with robust standard errors. The test of Wald was used in order to indicate the overall effect of explanatory variables. The result of this tests confirm that the explanatory variables including our control variables add significant explanatory power to the all the models from 1 to 6 (Appendix 3, Section 4.2). A categorical variable for the years from 2006 to 2010 (*n*=5 levels) is represented in *n-1* dummy variables, the base and omitted year is 2006.

**7.2.1 Independent Variables** On the one hand, Hypothesis 1a. proposes that the relative technological distances have an effect on the value of joint invention. Data revealed that both the relative technological distances, *TeachDist* ($\beta$=−0.9787,$p$<0.01) and *StuDist* ($\beta$=−0.6029,$p$<0.05), have an effect on the joint innovation value, thus confirming Hypothesis 1a. With all the controls in the model, the estimated coefficients for the relative technological distances in the learning dyad remain significant below the 5% level. On the other hand, Hypothesis 1b. states that the effect of the relative technological distances is negative on the joint invention value. According to data, Hypothesis 1b. is supported and thus indicates that overall, higher relative technological distances tend to lower the value of the joint invention of partner firms making them to be conservative and develop more proximate innovations to their profiles in the technological space. In other words, the closer the teacher and student firms in the learning dyad to their joint invention, the greater the likelihood of high value of the joint invention.

Regarding Hypothesis 2, the idea that the relative technological distance of the teacher firm has a higher negative effect on the value of joint invention than the relative technological distance of the student firm is tested. This hypothesis is confirmed in Model 6 with the relative technological distances variables being significant below the 5% level. The larger beta coefficient in the relative technological distance of the teacher firm, *TeachDist* ($\beta$=−0.9787,$p$<0.01), indicates a greater negative influence on the joint invention value than the student firm, *StuDist* ($\beta$=−0.6029,$p$<0.05). As previously discussed, this effect is presumed to occur because of the loose of ambidexterity in the alliance when at increasing the relative technological distance makes the teacher firm turn its exploitative quality into a more of the explorative type.

**Table 6. Negative Binomial Regression Models.**

| | ***Dependent variable**: JoinValue* | | | | | |
|---|---|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6[a] |
| ***Independent Variables*** | | | | | | |
| TeachDist | | -1.1954*** | -1.0340*** | -0.9799*** | -0.9787*** | -0.9787*** |
| | | (0.3041) | (0.3133) | (0.3124) | (0.3125) | (0.3166) |
| StuDist | | | -0.6851** | -0.7032*** | -0.6030** | -0.6029** |
| | | | (0.2753) | (0.2724) | (0.2701) | (0.2574) |
| TeachCap | | | | -0.0055*** | -0.0047** | -0.0046** |
| | | | | (0.0019) | (0.0019) | (0.0019) |
| StuCap | | | | | 0.0064*** | 0.0064*** |
| | | | | | (0.0022) | (0.0020) |
| ***Control Variables*** | | | | | | |
| GeoDist | 0.0336 | 0.0406 | 0.0350 | 0.0409 | 0.0464* | 0.0463* |
| | (0.0285) | (0.0281) | (0.0279) | (0.0278) | (0.0276) | (0.0226) |
| PriorColl | -0.0283*** | -0.0302*** | -0.0314*** | -0.0282*** | -0.0293*** | -0.0293*** |
| | (0.0045) | (0.0045) | (0.0046) | (0.0047) | (0.0048) | (0.0054) |
| PatCit | 0.0058** | 0.0059** | 0.0060** | 0.0064** | 0.0063** | 0.0063*** |
| | (0.0026) | (0.0026) | (0.0025) | (0.0025) | (0.0025) | (0.0022) |
| PCT | 0.4194** | 0.6088*** | 0.5854*** | 0.6088*** | 0.4979** | 0.4978** |
| | (0.2059) | (0.2095) | (0.2072) | (0.2059) | (0.2041) | (0.2109) |
| NumCod | 0.0155 | 0.0078 | 0.00004 | -0.0050 | -0.0093 | -0.0092 |
| | (0.0165) | (0.0163) | (0.0164) | (0.0163) | (0.0161) | (0.0158) |
| 2007 | -0.4056 | -0.3176 | -0.3563 | -0.3206 | -0.4680 | -0.4680 |
| | (0.2987) | (0.2933) | (0.2914) | (0.2917) | (0.2867) | (0.2739) |
| 2008 | -0.2531 | -0.0307 | -0.0137 | 0.0186 | -0.0518 | -0.0517 |
| | (0.2894) | (0.2862) | (0.2843) | (0.2835) | (0.2774) | (0.2911) |
| 2009 | -0.3113 | -0.1211 | -0.1042 | -0.1025 | -0.2463 | -0.2462 |
| | (0.2987) | (0.2932) | (0.2905) | (0.2898) | (0.2849) | (0.2800) |
| 2010 | -0.7729** | -0.7093** | -0.7066** | -0.6525** | -0.7161** | -0.7161** |
| | (0.3213) | (0.3156) | (0.3128) | (0.3107) | (0.3061) | (0.3009) |
| Constant | 0.5015 | 0.7106* | 0.9543** | 1.1273*** | 0.9271** | 0.9271** |
| | (0.3600) | (0.3679) | (0.3719) | (0.3755) | (0.3798) | (0.3860) |
| Log Likelihood | -697.8328 | -690.4526 | -687.2530 | -683.1363 | -679.1690 | -678.1689 |
| theta | 0.4469*** (0.0506) | 0.4837*** (0.0563) | 0.4999*** (0.0587) | 0.5156*** (0.0608) | 0.5382*** (0.0643) | 0.5381*** (0.0643) |
| AIC | 1,415.6660 | 1,402.9050 | 1,398.5060 | 1,392.2730 | 1,386.3380 | 1,386.338 |

*Note:* *p<0.1; **p<0.05; ***p<0.01; *Standard Errors in Parenthesis ; [a] Uses Robust Standard Errors.*

Hypothesis 3 is corroborated, since the effect of the technological capital of the student firm is positive and higher than the effect of the technological capital of the teacher firm. Concerning the effects, a positive and significant effect of the technological capital of the student firm, StuCap($\beta$=0.0064,$b$<0.01), and a negative impact of technological capital of the teacher firm, TeachCap($\beta$=−0.0046,$b$<0.05) is found. Therefore, it is tested that joint invention value is not only a function of the relative technological distances but also of the technological capital of the firms in the learning dyad context.

The results for technological capital are in line with Hypothesis 3. However, it was expected both coefficients to be positive, and according to the technological capital of the teacher firm that is not the case. A possible explanation for the negative effect might be that firms with wider technological capabilities, as the teacher firm in the learning dyad, have problems to apprehend knowledge that is technologically distant or highly radical because, as discussed, these organizations have an inclination to adapt solutions in existing knowledge areas (Ahuja & Lampert, 2001; Martin & Mitchell, 1998; Nooteboom et al., 2007). Likewise, some studies suggest that the exploitative role, of the teacher firm in this case, might tend to limit its explorative counterpart dimension (Kyriakopoulos & Moorman, 2004) making that going further into the explorative search, results in losing the ambidexterity balance in the alliance and consequently occasioning a low organizational performance for the firm and for the inter-organizational collaboration.

**7.2.2 Control variables** A total of five control variables were included in the negative binomial regression analyses of joint innovation value. Overall, only three were highly significant in Model 6 (PriorColl, PatCit, and PCT). On the one hand, the coefficients of patent citations, PatCit ($\beta$=0.0063,$p$<0.01), and Patent Cooperation Treaty, PCT ( $\beta$=0.4978,$p$<0.05), are all positives and significant, suggesting a positive relationship with the joint innovation value as expected.

On the other hand, contrary to other studies (Petruzzelli, 2011), prior collaboration, PriorColl ($\beta$=−0.0293,$p$<0.01) has a negative and significant effect on joint innovation value, also differing from the assumption stated in this study. A possible explanation may be related to the temporal effects on the knowledge bases of partner firms. In fact, it is recognized that innovations are the result of new combinations of existing knowledge (Nelson & Winter, 1982; Schumpeter, 1934). However, the more the firms collaborate together, the more their knowledge bases are becoming highly related through time which difficults the creation of new combinations of knowledge. Other possible reason is based on the learning opportunity of codified or tacit knowledge (Schildt et al., 2012). At their initial collaborations, the partner firms learn the knowledge elements that they find the easiest to absorb such as codified knowledge. Through time and after this early learning, they have to learn from areas that require greater ability and competencies in order to identify, assimilate, and absorb knowledge that has to be combined. The result is a more difficult creation of valuable innovations.

According to the data the variable *NumCod* is not significant. Thus, interestingly, and in line with recent research (Fischer & Leidinger, 2014) the number of IPC classes from which the joint patent is technologically characterized does not have any influence on the value of the joint invention. Contrary to what it was anticipated, geographical distance has a positive effect and has a slight significance in Model 6, *GeoDist* ($\beta$=0.0463,$p$<0.1). A possible explanation may be related to the need of having differences or diversity in terms of ideas and paradigms in order to develop valuable innovations. Although physical proximity in firms would favor interactions (Boschma, 2005), the apparent disadvantage of being in other country can be counterbalanced by strategies such as the use of Information Technologies, or the location of subsidiaries firms. Therefore, it seems that differences in cognitive paradigms resulting from the cultural differences should deserve more attention in order to improve the understanding of the positive effect of physical distance on the value of inventions. Results also indicate that a statistical difference with a negative and significant predictor is found in the year 2010. No statistical difference occurred with the other dummy year variables (2007, 2008, and 2009).

**7.2.3 Using other technological distance measures in the model** Although the complement of the cosine similarity or Jaffe's distance is the most used measure of technological distance in the literature it has some mathematical shortcomings. First, it is important to give a formal definition of distance: let $X$ be a set. A function $d{:}X \times X \rightarrow R$ is called a distance on $X$ if, for all $x,y \in X$, there holds (Deza & Deza, 2009):

i. $d(x,y) \geq 0$ (*non-negativity*)

ii. $d(x,y)=d(y,x)$ (*symmetry*)

iii. $d(x,x)=0$ (*reflexivity*)

According to this definition, Jaffe's measure is a distance and covers the notion as basic to human experience given the idea of measuring difference (O Sercoid, 2007).

However, an important standard for measurement of distance is the *metric*. Distance metrics and distances have become an essential tool in many areas and applications such as biology, code theory, engineering, clustering, networks, and other areas of science. Therefore, it is key to define what a *metric* is. A function $d{:}X \times X \rightarrow R$ is called a *metric* on $X$ if, and only if, for all $x,y,z \in X$, there holds (Deza & Deza, 2009):

i. $d(x,y) \geq 0$, with equality if, and only if, $x=y$ (*non-negativity*)

ii. $d(x,y)=d(y,x)$ (*symmetry*)

iii. $d(x,y) \leq d(x,z)+d(z,y)$ *(triangle inequality)*

In this event, the set $X$ is a *metric space* and, for each $x,y \in X$ the number $d(x,y)$ is the *distance* between $x$ and $y$ with respect to the metric $d$ (O Sercoid, 2007).

According to the above definition, the Jaffe's measure is not a metric because it does not hold the the i) and iii) properties.

In order to overcome these weaknesses of the Jaffe measure, the present research also used well-defined metric distances in order to measure the relative technological distances. First, the relative technological distance was measured using the well-known Euclidean distance. The results obtained (Appendix 3, Section 5.3) are in line with those of using the cosine similarity of Jaffe and confirmed the hypotheses. However, the independent variable of student relative technological distance is slightly less significant *StuDist* ($\beta$=−0.676248,$b$<0.1) than its counterpart when using Jaffe's measure. A possible explanation to this result could be that the Euclidean distance is a metric that does not hold the Independence of Irrelevant Patent Classes (IIPC).

The IIPC stablishes the following: for any R&D portfolio vectors $P0$,$P1$,$P2$, the relevant patent classes $k$ for $P0$ are those with $P0k$>0. "*A distance measure satisfies independence of irrelevant patent classes property when the distance between portfolio* P0 *and each of the two technology profiles* $P1$,$P2$ *is the same whenever these two portfolios differ from each other only in irrelevant patent classes, but have equal shares in all relevant patent classes*" (Bar & Leiponen, 2012, pp. 458). In order to have a metric that satisfies the independence of irrelevant patent classes, the present research used also the Min-complement as a measure of the relative technological distances. The results obtained using the Min-complement metric confirmed the hypotheses again (Appendix 3, Section 5.2). Moreover, the significance of the coefficients of independent and control variables are in line to the models using Jaffe's distance measures.

**7.2.4 Summary of key results** Table 7 has a summary of all the hypothesis and the corresponding results. Overall, the results provide evidence that both relative technological distances affect the joint invention value of firms in the learning dyad.

Two hypothesis (H1a., H1b., and H2) presented in the model were supported and significant. Likewise, results also present evidence about the effect on joint invention value not only from the relative technological distance, but also from the technological capital of both firms in the learning dyad (Hypothesis 3).

It was found that the relative technological distances have a negative effect in joint invention value (Hypothesis 1a. and 1b.). To summarize, this result is saying that the farther the relative technological distance is, the less likely to have a higher value of the join invention in the learning dyad. In addition, it was found that the relative technological distance of the student firm has lower negative effect than the teacher firm in the value of joint invention (Hypothesis 2). Therefore, the data provides the message that increasing the teacher's relative technological distance has a more pronounced negative effect than increasing the student's relative technological distance on joint invention value.

**Table 7. Summary of key results**

| Hypothesis | Proposed Sign | Finding |
|---|---|---|
| H1a. and H1b. Relative technological distances | Negative | Negative, Significant ($p<0.05$) |
| H2. Teacher's relative technological distance have a stronger and negative influence than student's relative technological distance | Larger beta on teacher's relative technological distance coefficient | Supported |
| H3. Student's technological capital has a stronger and positive effect than teacher's technological capital | Positive and larger beta on student's technological capital coefficient | Partially supported. Teacher's technological capital coefficient is negative. |

Turning now to scope of technological capital, the results of the study also provide support for the hypothesis that technological capital of the student firm has a positive and higher effect on join invention value than the technological capital of the teacher firm (Hypothesis 3). The unexpected result of the study, that technological capital of the teacher firm has a negative effect on join invention value, deserves more attention. To summarize, this result is saying that the greater the technological capital of the teacher firm is, the lower the likelihood is that the joint invention value has a higher value. A possible explanation for this interesting result can have two parts. The first one is that firms with broad knowledge bases have no inclinations for non-related knowledge or technology. The second one is that, in order to maximize the learning dyad performance via ambidexterity, there is a limitation of the teacher firm for explorative capabilities taking into account its exploitative role.

# 8. CONCLUSIONS

Based on the empirical contributions presented in the research, this last chapter has been divided in three sections. First, the dissertation's contributions to theory and research are discussed. The second section describes the contribution to practice. Last, some directions for future research are presented, concluding the dissertation.

## 8.1 CONTRIBUTION TO THEORY AND RESEARCH

Research on innovation theory has either pursued the search of opportunities that individual firms find beyond its knowledge boundaries to innovate, or the innovation performance of inter-organizational collaborations such as strategic alliances. Research examining the combination of innovation search of opportunities beyond firms' knowledge limits in technological space in a context of strategic alliances has been scarce in the literature. The contributions of the present research to this paradigm is clarified from three related standing points explained next: the inter-organizational collaborations, the relative absorptive capacity, and the Knowledge Based View of the firm.

From the inter-organizational collaborations theory, the present research shed further light on exploring how technological attributes may influence innovation performance. The study contributes to the existing literature on firm alliances extending previous research in two main directions. First, the study combines the technological attributes of firms, controlling for relational attributes such as prior collaborations and geographical distance, with the impact of the joint innovation location in the technological space as a strategy for the creation of value. Existing literature on technological distance in alliances mostly focuses on the distance between firms, assuming that alliances with certain technological distance between firms would have equally innovative performance. This argumentation is neglected in the present research by focusing the attention on the technological position of the joint invention as part of the alliance's overall technology strategy. This emphasis on the joint invention originates two distances to each of the knowledge bases of firms in the alliance, the relative technological distances, a topic seldom addressed in previous research. Therefore, an important contribution of this study is its emphasis on *how* the potential value of jointly developed technological innovations in alliances is influenced by differences across technological domains and specializations reflected on these relative technological distances.

Second, the present research examines why some alliances are better than others showing under which circumstances collaborations are more valuable in terms of joint innovations. The rationale from this study is motivated by the focus of firms on developing technology located in a space outside their own technological boundaries. The research suggests that an inappropriate stress on overcoming the limits of local opportunities searches by renovating through relative technological remote knowledge is unlikely to maximize the value of their joint invention. Too much relative technological distance may be detrimental since the development of valuable innovations need to balance the high level of uncertainty presented in alliances through a more secure strategy via the development of technologies in the related knowledge domains of firms. Therefore, successfully collaborating firms need to be, in general, in technological proximity to their joint invention in order to generate value.

From the relative absorptive capacity perspective, the contribution of this research is based on two conceptual premises: First, firms in the alliance have not only to identify valuable external knowledge, but need to counterbalance its ability to assimilate it, and principally apply it in order to create value (Cohen & Levinthal, 1990). And second, this learning capacity needs to account for knowledge or technological relative characteristics from firms in the the learning dyad (Lane & Lubatkin, 1996). According to Lane & Lubatkin, firms in the learning dyad have roles of teacher and student in their interactive learning processes. Although their study does not make explicit any clear conceptual or practical delimitation in order to differentiate between teacher or student beyond the exemplification by industrial domains in biotechnology and pharmaceutical, the present research contributes making that differentiation between teacher and student firms based on comparing how broad is the knowledge and expertise of the firms. Even though, this first approximation can appear simple to the eyes of the technology transfer and organizational learning theories, it is an initial approach that makes a clear difference and has important implications as shown by the results obtained from the attributes of teacher and student firms. Thus, the examination of the learning dimension of the present research draws on the mentioned relative characteristics of firms originating a new view of two important determinants on innovation value: teacher and student technological distances and teacher and student technological capitals.

The notion of technological distance is based on how far apart or close are knowledge bases in the technological space. Consequently, taking into account the differentiated roles firms have in the learning dyad and the joint invention, the research analysis draws on the idea that firms can undertake two types of technological opportunity search beyond their knowledge boundaries: local search (or exploitative), and a more distant search (explorative). From this perspective, it is suggested that long technological distance reflects more exploratory search of innovative opportunities, while a short distance reflects the outcome of more exploitative search. As mentioned, the role of firms originates what is called in this research "*relative technological distances*". These distances are relative to the potential innovative opportunity that, in the context of alliances, is represented by the joint innovation developed by firms in the learning dyad.

Therefore, the contribution made by the present research on differentiating explicitly between teacher and student firms in the learning dyad have two implications: First, it makes research goes further from the technological distance between firms concept offered by most of literature on relative absorptive capacity of knowledge, and breaks it into relative technological distances focusing on the joint invention as a transitivity object through which the distance of firms are relatively located in technological space. Second, the technological capital is also affected by this learning dyad role decomposition differentiating between a teacher firm technological capital, and a student firm technological capital. Results show a dependence of the technological capital of firms on the role of the firm in the learning dyad. Specifically, there might be limitations of expanding the technological capital of the teacher firm in order to obtain high valuable innovations, something that does not occur with its counterpart the student firm. This postulates a counterintuitive idea: not always having a high technological capital by a firm in the alliance indicates more probability to obtain high valuable innovations. This behavior of the technological capital of the teacher firm might be presumably a result of the great effort needed to go farther from its technological boundaries, and also because the potential loss of the ambidexterity balance in the alliance at higher technological distances which bring low organizational performance. In summary, this is a new form to analyze technological characteristics of firms on alliances from the relative absorptive capacity perspective which reinforces the statements mentioned previously: not only the firms' technological capital, and knowledge bases, but the location of the joint invention in technological space is a critical element to generate value.

Finally, from the Knowledge Based View –KBV– of the firm, were knowledge is the most strategic resource (Grant, 1996), the present research has three important elements that help to improve the understanding of collaboration and innovation: First, this research clarifies the existence of firms as a consequence of their ability to create knowledge (Kogut & Zander, 1992; Nonaka & Takeuchi, 1995) and makes effective use of it (Rebolledo & Nollet, 2011) by arguing that knowledge characteristics are critical to indicate certain directions of success or failure to create valuable knowledge. Second, the study goes further from the effort of merely identifying knowledge as the basis of competitive advantage, but into explaining how organizations can develop it (Argote, Ingram, Levine, & Moreland, 2000) by examining the relative technological attributes of firms in order to trace how firms' knowledge mechanisms support the development of joint innovations. Finally, not only the identification but the differentiation of firms in the learning dyad setting were the knowledge attributes are structured, helps to understand how collaboration efficiently exploits and incorporates knowledge. In sum, the overall contribution of this research is based on the better understanding of what the KBV in the context of alliances calls donor and recipient firms characteristics for the creation of competitive advantages (Grant, 1996); centering the attention not only on alliance partners, but on the created knowledge.

## 8.2 CONTRIBUTIONS TO PRACTICE

In terms of managerial practical implications, the findings deliver useful insights for firms engaged in the development of new innovative solutions. Once assumed the accompanying economic risks and rewards of alliances, elucidating the contributing factors of innovation value is of great significance to the managers and practitioners. Specifically, the research revealed that the joint development of innovation in an alliance appears to be led by three factors: (i) the differentiation of the firms as a student or as a teacher in the context of the '*learning dyad*', (ii) the location of the joint invention in the technological space, and (iii) the counterbalance effect of the technological capital of teacher and student firms forming the alliance. In so doing, these three elements should be taken carefully into account in a partner selection process *ex ante* the joint innovation development.

Regarding the learning dyad configuration, the research clearly indicates the differences between the teacher and student firms of the alliance. Therefore, firms in alliances not only have to be distinguished by their size archetypally found in the business environment as large or small firms; but also as organizations that have specific roles as exploitative –teacher firm– and explorative –student firm– knowledge features. In so doing, managers should be aware that alliances may benefit from both the existence of a broad knowledge embedded in the teacher firm which favor the common language setting and the transfer of knowledge, and a more narrow or dissimilar competence embedded in the student firm, which is fundamental for the combination and creation of Schumpeterian innovations. Furthermore, technological capital has also a dissimilar relative effect depending on the firm's role as a teacher or as a student. Thus, managers and practitioners may be able to identify firm roles in order to take action about the technological capital depending on the positive or negative effects of being a student or teacher firm to obtain high valuable innovations.

Additionally, the results show the negative effects or difficulties that alliances can have in going to innovate in extremely unexplored technological areas. Therefore, the importance of having a vision in advance of the potential joint innovation to be developed by the alliance, might help to overcome the disadvantage of excessive relative technological distances outside the firms' technological domains. This offers a cautionary message to alliance's managers, suggesting a selective and restricted use of the innovative step that firms can take to obtain a successful joint technology.

The use of patents in the present study can help managers operationalize into actionable items the research results obtained in order to make decisions. Therefore, the view of a competitive intelligent system via patent search and analysis, is not an unreasonable idea that might be used as a tool that support the decision making process. Accordingly, this research sheds light about strategies of how to shape joint technological knowledge maximizing the alliance’s chances of creating high value and achieve competitive advantages.

In terms of policy making, the above arguments indicate that the promotion and establishments of firms’ alliances and the funding of collaborative R&D projects should consider partner and context characteristics. Accordingly, policies should take into account not only financial and economic components, but technological knowledge characteristics of firms and the potential innovation. Policy makers are suggested to appreciate that such attributes may significantly affect collaborations and need to be carefully analyzed when discussing the potential results. For that reason, policies might stimulate alliances in a selective form preferring collaborations between firms that pretend to develop not too far technological innovations.

Finally, the results questions the importance of technological capital from the firms for facilitating knowledge transfer and creation. In particular, contrary to the intuitive thought of increasing technological capital to generate more innovations, it emerges that successful firms’ collaboration tend to occur when the teacher partner does not have an extremely high technological capital. In this respect, government policy should promote a matching between firms that enable them to choose the best partner in terms of knowledge attributes for a determined innovation creation, taking into account that not in all scenarios the firm with higher knowledge capital is always the best option.

## 8.3 DIRECTION FOR FUTURE WORK

Finally, there is a number of limitations of the study that may embody a venue for future research. First, the sample was restricted to firms in the biotechnology industry and so the results may not be generalized to firms in other industries. Future research should include other industries besides the biotechnology sector with a potential creation of joint technological knowledge. Second, the sample was also restricted to the United States Patent and Trade Office database taking into account the differences on the intellectual property processes followed when compared to the other offices such as the European Patent Office or the Japan Patent Office. Thus, extending the studies to include intellectual management offices in other latitudes in order to find differences or to validate the result are of such importance. Third, another limitation is related to the fact that patents measure partially the production of innovations and are a reflection of codified knowledge. Future work can consider other types of intellectual property artefacts besides patents, and even go further exploring in more detail the effects exerted by tacit knowledge embedded in organizational routines on the creation of valuable innovations. Fourth, the present research includes only firms. Including variety of organizations (e.g. universities, governmental organizations, foundations, public or private organizations) and even individual people will help to find critical differences and similarities between actors. This will help to understand the specific mechanisms used to generate value depending of the type of organizations or actors. Fifth, because the study is centered on dyadic alliances, future research should go further at the network level in order to include more complex inter-organizational collaborations and analyze the effects of these kinds of interactions on innovation value.

From the absorptive capacity perspective, the study is centered on the technological dimension of firms. Future work should include distances related to the knowledge embedded in the organizational and dominant logics of firms (Lane & Lubatkin, 1996) as learning factors that may affect also the value of innovations. As mentioned, the differentiation between teacher firm and student firm was made based on the knowledge and expertise attributes of firms. More accurate differentiation can be studied in order to include other dimensions of the technology transfer, organizational learning and innovation theory. Likewise, new studies should embrace new contexts in terms of the roles of firms in alliances besides the '*learning dyad*'. In this line, firms in the alliances should have complementary roles defined from the framework of the two types of absorptive capacity: the potential absorptive capacity, and the realized absorptive capacity (Zahra & George, 2002). Finally, the negative effect of technological capital will deserve a special treatment. Future possible paths comprise the study of the desorptive capacity which focus on the source over the recipient of knowledge in the technology transfer process, something rarely found in the academic literature (Lichtenthaler & Lichtenthaler, 2010), and organizational unlearning and relearning (Zhao, Lu, & Wang, 2013) to analyze respectively the discarding of the outdated and useless knowledge and the dynamic acquisition of new knowledge that may affect the behavior of technological capital of the teacher firm.

# BIBLIOGRAPHY


ACOSTA, M., CORONADO, D., & MARTINEZ, M. A. Spatial differences in the quality of university patenting: Do regions matter? *Research Policy*, 2012 *41*(4), 692–703.

AGHION, P., BLOOM, N., BLUNDELL, R., GRIFFITH, R., & HOWITT, P. Competition and innovation: an inverted-u relationship. *The Quarterly Journal of Economics*, 2005 *120*(2), 701–728.

AHUJA, G. Collaboration Networks , Structural Holes , and Innovation : A Longitudinal Study. *Administrative Science Quarterly*, 2000 *45*(3), 425–455.

AHUJA, G., & KATILA, R. Technological acquisitions and the innovation performance of acquiring firms: A longitudinal study. *Strategic Management Journal*, 2001 *22*(3), 197–220.

AHUJA, G., & KATILA, R. Where do resources come from? The role of idiosyncratic situations. *Strategic Management Journal*, 2004 *25*(89), 887–907.

AHUJA, G., & LAMPERT, C. M. Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, 2001 *22*(6-7), 521–543.

AKERS, N. J. The referencing of prior art documents in European patents and applications. *World Patent Information*, 2000 *22*(4), 309–315.

ALBERT, M. B., AVERY, D., NARIN, F., & McALLISTER, P. Direct validatin of citation couts as indicators of industrially important patents. *Research Policy*, 1991 *20*(3), 251–259.

ALCÁCER, J., & GITTELMAN, M. Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations. *Review of Economics and Statistics*, 2006 *88*(4), 774–779.

ALMEIDA, P., & KOGUT, B. Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science*, 1999 *45*(7), 905–917.

ANDRIOPOULOS, C., & LEWIS, M. W. Exploitation-Exploration Tensions and Organizational Ambidexterity: Managing Paradoxes of Innovation. *Organization Science*, 2009 *20*(4), 696–717.

ANTONELLI, C. Collective Knowledge Communication and Innovation: The Evidence of Technological Districts. *Regional Studies*, 2000 *34*(6), 37–41.

ANTONELLI, C. Localized product innovation: the role of proximity in the Lancastrian product space. *Information Economics and Policy*, 2004 *16*(2), 255–274.

AOKI, R., & HU, J. L. Licensing vs. Litigation: The Effect of the Legal System on Incentives to Innovate. *Journal of Economics & Management Strategy*, 1999 *8*(1), 133–160.

ARCHIBUGUI, D., & SIRILLI, G. The Direct Measurement of Technological Innovation in Business. In *International Conference on Innovation and Enterprise: Statistics and Indicators* 2001 (pp. 2–47). Rome.

ARGOTE, L. *Organizational learning: Creating, retaining, and transfering knowledge* (2nd ed., Vol. 1). New York: Springer Heidelberg. 2013

ARGOTE, L., INGRAM, P., LEVINE, J. M., & MORELAND, R. L. Knowledge Transfer in Organizations: Learning from the Experience of Others. *Organizational Behavior and Human Decision Processes*, 2000 *82*(1), 1–8.

ARGOTE, L., MCEVILY, B., & REAGANS, R. Managing knowledge in organizations: An integrative framework and review of emerging themes. *Management Science*, 2003 *49*(4), 571–582.

ASLESEN, H. W., & JAKOBSEN, S. The role of proximity and knowledge interaction between head offices and KIBS. *Tijdschrift Voor Economische En Sociale Geografie*, 2007 *98*(2), 188– 201.

BAR, T., & LEIPONEN, A. A measure of technological distance. *Economics Letters*, 2012 *116*(3), 457–459.

BARNEY, J. Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 1991 *17*(1), 99–120.

BARON, J., & DELCAMP, H. The private and social value of patents in discrete and cumulative innovation. *Scientometrics*, 2012 *90*(2), 581–606.

BELDERBOS, R., CASSIMAN, B., FAEMS, D., LETEN, B., & VAN LOOY, B. Co-ownership of intellectual property: Exploring the value-appropriation and value-creation implications of co-patenting with different partners. *Research Policy*, 2014 *43*(5), 841–852.

BELSLEY, D. A. *Conditional diagnostics: collinearity and weak data in regression*. Hoboken, New Jersey: Jhon Wiley & Sons. 1991

BELSLEY, D. A., KUH, E., & WELSCH, R. E. *Regression diagnostics: identifying influential data and sources of collinearity* (1st ed.). Hoboken, New Jersey: Jhon Wiley & Sons. 2004

BENNER, M. J., & TUSHMAN, M. L. Exploitation, exploration, and process management: The productivity dilemma revisited. *Academy of Management Review*, 2003 *28*(2), 238–256.

BENNER, M., & WALDFOGEL, J. Close to you? Bias and precision in patent-based measures of technological proximity. *Research Policy*, 2008 *37*(9), 1556–1567.

BEUZEKOM, B. V., & ARUNDEL, A. Biotechnology Patents. In *OECD Biotechnology Statistics 2009* (p. 103). Paris: OECD Publishing. 2009

BIERLY, P. E., DAMANPOUR, F., & SANTORO, M. D. The application of external knowledge: Organizational conditions for exploration and exploitation. *Journal of Management Studies*, 2009 *46*(3), 481–509.

BLOOM, N., SHANKERMAN, M., & VAN REENEN, J. Identifying Technology Spillovers and Product Market Rivalry. *Econometrica*, 2013 *81*(4), 1347–1393.

BÖNTE, W., & WIETHAUS, L. Knowledge disclosure and transmission in buyer-supplier relationships. *Review of Industrial Organization*, 2007 *31*(4), 275–288.

BONTIS, N. Intellectual capital : an exploratory study that develops measures and models. *Management Decision*, 1998 *36*(2), 63–76.

BOOTH, R. The measurement of intellectual capital. *Management Accounting*, 1998 *76*, 26–28.

BORGATTI, S. P., EVERETT, M. G., & FREEMAN, L. C. Ucinet. Analytic Technologies. 1999

BORGATTI, S. P., EVERETT, M. G., & JHONSON, J. C. *Analyzing social networks*. London: SAGE Publications. 2013

BOSCHMA, R. Proximity and Innovation: A Critical Assessment. *Regional Studies*, 2005 *39*(1), 61–74.

BOSCHMA, R., & FRENKEN, K. The spatial evolution of innovation networks. A proximity perspective. In R. Boschma & R. Martin (Eds.), *The Handbook of evolutionary economic geography* (pp. 120–136). Cheltenham, UK: Edward Elgar Publishing Inc. 2010

BRADFORD, W., & RAJAT, K. Using Patent Information In Technology Business Planning-I. *Research Technology Management*, 1989 *32*(1), 36–42.

BRADLEY, K. Intellectual capital and the new wealth of nations. *Business Strategy Review*, 1997 *8*(1), 53–62.

BRANSTETTER, L. G. Are knowledge spillovers international or intranational in scope? Microeconometric evidence from the U.S. and Japan. *Journal of International Economics*, 2001 *53*(1), 53–79.

BRESCHI, S., LISSONI, F., & MALERBA, F. Knowledge-relatedness in firm technological diversification. *Research Policy*, 2003 *32*(1), 69–87. Retrieved from

BRIGGS, K. Co-owner relationships conducive to high quality joint patents. *Research Policy*, 2015 *44*(8), 1566–1573.

BRIGGS, K., & WADE, M. More is better: evidence that joint patenting leads to quality innovation. *Applied Economics*, 2014 *46*(35), 4370–4379.

BROEKEL, T., & BOSCHMA, R. Knowledge networks in the Dutch aviation industry: The proximity paradox. *Journal of Economic Geography*, 2012 *12*(2), 409–433.

BUERGER, M., & CANTNER, U. The regional dimension of sectoral innovativeness: An empirical investigation of two specialized suppliers and two science-based industries. *Papers in Regional Science*, 2011 *90*(2, SI), 373–393.

CALABRESE, T., BAUM, J. A., & SILVERMAN, B. S. Canadian biotechnology start-ups, 1991-1997: the role of incumbents' patents and strategic alliances in controlling competition. *Social Science Research*, 2000 *29*(4), 503–534.

CALOGHIROU, Y., KASTELLI, I., & TSAKANIKAS, A. Internal capabilities and external knowledge sources: complements or substitutes for innovative performance? *Technovation*, 2004 *24*(1), 29–39.

CAMERON, A. C., & TRIVEDI, P. K. Regression-based tests for overdispersion in the Poisson model. *Journal of Econometrics*, 1990 *46*(3), 347–364.

CAMERON, A. C., & TRIVEDI, P. K. *Regression analysis of count data* (2nd ed.). New York: Cambridge University Press. 2013

CAMISÓN, C., & FORÉS, B. Knowledge absorptive capacity: New insights for its conceptualization and measurement. *Journal of Business Research*, 2010 *63*(7), 707–715.
CANTNER, U., & GRAF, H. The network of innovators in Jena: An application of social network analysis. *Research Policy*, 2006 *35*(4), 463–480.

CANTWELL, J. A., & BARRERA, P. The localisation of corporate technological trajectories in the interwar cartels: Cooperative learning versus an exchange of knowledge. *Economics of Innovation & New Technology*, 1998 *6*(2/3), 257.

CANTWELL, J., & COLOMBO, M. G. Technological and Output Complementarities: Inter-Firm Cooperation in Information Technology Ventures. *Journal of Management and Governance*, 2000 *4*(1), 117–147.

CAO, Q., GEDAJLOVIC, E., & ZHANG, H. Unpacking Organizational Ambidexterity: Dimensions, Contingencies, and Synergistic Effects. *Organization Science*, 2009 *20*(4), 781–796.

CARSON, E., RANZJIN, R., WINEFIELD, A., & MARSDEN, H. Intellectual capital: mapping employee and work group attributes'. *Journal of Intellectual Capital*, 2004 *5*(3), 443–463.

CASCIARO, T. Determinants of Governance Structure in Alliances: the Role of Task, Strategic, and Partner Uncertainties. *Industrial and Corporate Change*, 2003 *12*(6), 1223–1251.

CASSIMAN, B., COLOMBO, M. G., GARRONE, P., & VEUGELERS, R. The impact of M&A on the R&D process - An empirical analysis of the role of technological- and market-relatedness. *Research Policy*, 2005 *34*(2), 195–220.

CHESBROUGH, H. *Open Innovation*. Boston, Massahusetts: Harvard Business School Press. 2003

COHEN, B. H. *Explaining Psychological Statistics* (4th ed.). Hoboken, New Jersey: Jhon Wiley & Sons Ltda. 2013

COHEN, J., COHEN, P., WEST, S., & AIKEN, L. *Applied Multiple Regression / Correlation Analysis for the Behavioral Sciences* (3rd ed.). London: Lawrence Erlbaum Associates. 2003

COHEN, W. M., & LEVINTHAL, D. A. Innovation and learning: the two faces of R & D. *The Economic Journal*, 1989 *99*(397), 569–596.

COHEN, W. M., & LEVINTHAL, D. A. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 1990 *35*(1), 128–152.

COHEN, W. M., & LEVINTHAL, D. A. Fortune favors the prepared firm. *Management Science*, 1994*40*(2), 227–251.

COLOMBO, M. G. Alliance form: A test of the contractual and competence perspectives. *Strategic Management Journal*, 2003 *24*(12), 1209–1229.

COMTE, A. *The Positive Philosophy of Auguste Compte*. (H. Martineau, Ed.) (Vol. 1). London: George Bell & Sons. 1896

CORNELL, U., INSEAD, & WIPO. *The Global Innovation Index 2015*. (S. Dutta, B. Lanvin, & S. Wunsch-Vincent, Eds.) *Effective Innovation Policies for Development*. Geneva, Switzerland: World Intellectual Property Organization. 2015

COWAN, R., & JONARD, N. If the Alliance Fits ...: Innovation and Network Dynamics. In *Network Strategy: Advances in Strategic Management* (1st ed., Vol. 25, pp. 427–455). Oxford: Emerald Group Publishing Limited. 2008

CRESWELL, J. W. *Research Design: qualitative, quantitative and mixed methods approach* (4th ed.). California, USA: SAGE Publications. 2014

CRISCUOLO, P., & VERSPAGEN, B. Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research Policy*, 2008 *37*(10), 1892–1908.

CUNNINGHAM, S. W., & WERKER, C. Proximity and collaboration in European nanotechnology. *Papers in Regional Science*, 2012 *91*(4), 723–742.

DANE. *Encuesta de Desarrollo e Innovación Tecnológica Industria Manufacturera - EDIT VII*. Bogotá. 2015

DANGELICO, R. M., GARAVELLI, A. C., & PETRUZZELLI, A. M. A system dynamics model to analyze technology districts' evolution in a knowledge-based perspective. *Technovation*, 2010 *30*(2), 142–153.

DAS, T. K., & RAHMAN, N. Partner Misbehaviour in Strategic Alliances - Guidelines for Effective Deterrence. *Journal of General Management*, 2001 *27*(1), 43–70.

DAVIS, P. S., ROBINSON JR, R. B., PEARCE II, J. A., & PARK, S. H. Business unit relatedness and performance: A look at the pulp and paper industry. *Strategic Management Journal*, 1992 *13*(5), 349–361.

DE CLERCQ, D., & SAPIENZA, H. J. When do venture capital firms learn from their portfolio companies? *Entrepreneurship: Theory and Practice*, 2005 *29*(4), 517–535.

DE JONG, J. P. J., & FREEL, M. Absorptive capacity and the reach of collaboration in high technology small firms. *Research Policy*, 2010 *39*(1), 47–54.
DECAROLIS, D., & DEEDS, D. The impact of stocks and flows of organizational knowledge on firm performance. An empirical investigation of the biotechnology industry. *Strategic Management Journal*, 1999 *20*(10), 953–968.

DENZIN, N. K., & LINCOLN, Y. *Handbook of qualitative research*. (N. K. Denzin & Y. S. Lincoln, Eds.) (3rd ed.). Lonodn: Sage Publications. 2005

DENZIN, N. K., LINCOLN, Y. S., & GUBA, E. G. Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (4th ed., pp. 97–128). Thousand Oaks: SAGE Publications. 2011

DEZA, M. M., & DEZA, E. *Encyclopedia of Distances*. *PhD Proposal* (Vol. 1). London, New York: Springer Berlin Heidelberg. 2009

DICKSON, P. H., & WEAVER, K. M. Environmental determinants and individual-level moderators of alliance use. *Academy of Management Journal*, 1997 *40*(2), 404–425.

DOZ, Y. L. Technology Partnerships between Larger and Smaller Firms: Some Critical Issues. *International Studies of Management & Organization*, 1988 *17*(4), 31–57.

DUYSTERS, G., & DE MAN, A. P. Transitory alliances: an instrument for surviving turbulent industries? *R & D Management*, 2003 *33*(1), 49–58.

DYER, J. H. The determinants of trust in supplier-automaker relationships in the U.S, Japan, and Korea. *Journal of International Business Studies*, 2002 *31*(2), 259–285.

DYER, J. H., & CHU, W. The Role of Trustworthiness in Reducing Transaction Costs and Improving Performance: Empirical Evidence from the United States, Japan, and Korea. *Organization Science*, 2003 *14*(1), 57–68.

DYER, J. H., & HATCH, N. W. Relation-specific capabilities and barriers to knowledge transfers: Creating advantage through network relationships. *Strategic Management Journal*, 2006 *27*(8), 701–719.

DYER, J. H., & SINGH, H. The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review*, 1998 *23*(4), 660–679.

EASTERBY-SMITH, M., LYLES, M. A., & TSANG, E. W. K. Inter-organizational knowledge transfer: Current themes and future prospects. *Journal of Management Studies*, 2008 *45*(4), 677–690.

EASTERBY-SMITH, M., THORPE, R., & JACKSON, P. R. *Management Research* (3rd ed.). London: SAGE Publications Ltd. 2008

EDVINSSON, L., & SULLIVAN, P. Developing a model for managing intellectual capital. *European Management Journal*, 1996 *14*(4), 356–364.

EISENHARDT, K. M., & MARTIN, J. E. Dynamic Capabilities: What are they? *Strategic Management Journal*, 2000 *21*(1), 1105–1121.

ENKEL, E., & GASSMANN, O. Creative imitation: exploring the case of cross-industry innovation. *R & D Management*, 2010 *40*(3), 256–270.

ERNST, H. Patent applications and subsequent changes of performance: Evidence from time-series cross-section analyses on the firm level. *Research Policy*, 2001 *30*(1), 143–157.

FARJOUN, M. The independent and joint effects of the skill and physical bases of relatedness in diversification. *Strategic Management Journal*, 1998 *19*(7), 611–630.

FISCHER, T., & LEIDINGER, J. Testing patent value indicators on directly observed patent value - An empirical analysis of Ocean Tomo patent auctions. *Research Policy*, 2014 *43*(3), 519–529.

FLAVIO, A., & AVILA, D. Total Factor Productivity Growth in Agriculture: The Role of Technological Capital. In *Handbook of Agricultural Economics* (1st ed., Vol. 4, pp. 3769–3822). Burlington: Academic Press. 2010

FLEMING, L. Recombinant Uncertainty in Technological Search. *Management Science*, 2001 *47*(1), 117–132.

FONTANA, R., & NESTA, L. Product Innovation and Survival in a High-Tech Industry. *Review of Industrial Organization*, 2009 *34*(4), 287–306.

FORNAHL, D., BROEKEL, T., & BOSCHMA, R. What drives patent performance of German biotech firms? The impact of R&D subsidies, knowledge networks and their location. *Papers In Regional Science*, 2011 *90*(2, SI), 395–418.

FUNG, M. . K., & CHOW, W. . W.  Measuring the intensity of knowledge flow with patent statistics. *Economics Letters*, 2002 *74*(3), 353–358.

GALUNIC, C., & RODAN, S. Resource recombinations in the firm: knowledge structures and the potential for Schumpeterian innovation. *Strategic Management Journal*, 1998 *19*(12), 1193–1201.

GAMBARDELLA, A., HARHOFF, D., & VERSPAGEN, B. The value of European patents. *European Management Review*, 2008 *5*, 69–84.

GARCIA, R., CALANTONE, R., & LEVINE, R. The Role of Knowledge in Resource Allocation to Exploration versus Exploitation in Technologically Oriented Organizations. *Decision Sciences*, 2003 *34*(2), 323–349.

GATIGNON, H., TUSHMAN, M. L., SMITH, W., & ANDERSON, P. A structural approach to assesing innovations: Construct development of innovation locus, type, and characteristics. *Management Science*, 2002 *48*(9), 1103–1122.

GAY, C., & LE BAS, C. Uses without too many abuses of patent citations or the simple economics of patent citations as a measure of value and flows of knowledge. *Economics of Innovation and New Technology*, 2005 *14*(5), 333–338.

GIBSON, C. B., & BIRKINSHAW, J. Antecedents, consequences, and mediating: Role of organizational ambidexterity. *Academy of Management Journal*, 2004 *47*(2), 209–226.

GILSING, V., NOOTEBOMM, B., VANHAVERBEKE, W., DUYSTERS, G., & VAN DER OORD, A. Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research Policy*, 2008 *37*(10), 1717–1731.

GRANT, R. M. Toward a Knowledge-Based Theory of the Firm. *Strategic Management Journal*, 1996 *17*, 109–122.
GRANT, R. M. *Contemporary Strategy Analysis* (8th ed.). Chichester, United Kingdom: Jhon Wiley & Sons Ltda. 2013a

GRANT, R. M. Reflections on knowledge-based approaches to the organization of production. *Journal of Management and Governance*, 2013b *17*(3), 541–558.

GRANT, R. M., & BADEN-FULLER, C. A Knowledge-Based Theory of Inter-Firm Collaboration. In *Academy of Management Proceedings* (pp. 17–21). 1995

GRANT, R. M., & BADEN-FULLER, C. A knowledge accessing theory of strategic alliances. *Journal of Management Studies*, 2004 *41*(1), 61–84.

GREEN, J. R., & SCOTCHMER, S. On the division of profit in sequential innovation. *The RAND Journal of Economics*, 1995 *26*(1), 20–33.

GREEVEN, M., & XIAODONG, Z. Innovation , Competences and the Role of Knowledge Networks in Hangzhou ' s Software Industry. *The Network Xperience*, 2009 *13*(1), 193–209.

GREUNZ, L. Geographically and technologically mediated knowledge spillovers between European regions. *Annals of Regional Science*, 2003 *37*(4), 657–680.

GRILICHES, Z. Issues in assessing the contribution and development of research to productivity growth. *The Bell Journal of Economics*, 1979a *10*(1), 92–116.

GRILICHES, Z. Issues in Assessing the Contribution of Research and Development to Productivity Growth. *The Bell Journal of Economics*, 1979b *10*(1), 92 – 116.

GRILICHES, Z. Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 1990 *27*(4), 1661–1707.

GRILICHES, Z. The Search for R&D Spillovers. *Scandinavian Journal of Economics*, *94*(Supplement), 1992 29–47.

GRUPP, H., & SCHMOCH, U. Perceptions of Scientification of Innovation as Measured by Referencing Between Patents and Papers: Dynamics in Science-Based Fields of Technology. In H. Grupp (Ed.), *Dynamics of Science-Based Innovation* (pp. 73–128). Berlin, Heidelberg: Springer Berlin Heidelberg. 1992

GUBA, E. G. The alternative paradigm dialog. In E. G. Guba (Ed.), *The paradigm dialog* (pp. 17–30). London: Sage Publications. 1990

GUELLEC, D., & VAN POTTELSBERGHE DE LA POTTERIE, B. The Value of Patents and Patenting Strategies: Countries and Technology Areas Patterns. *Economics of Innovation and New Technology*, 2002 *11*(2), 133–148.

GULATI, R. Does Familiarity Breed Trust ? the Implications of Repeated Ties for. *Academy of Management*, 1995a *38*(1), 85–112.

GULATI, R. Social Structure and Alliance Formation Patterns : A Longitudinal Analysis. *Administrative Science Quarterly*, 1995b *40*(4), 619–652.

GUPTA, A. K., SMITH, K. G., & SHALLEY, C. E. The interplay between exploration and exploitation. *Academy of Management Journal*, 2006 *49*(4), 693–706.

HAGEDOORN, J. the Rationale of Strategic Understanding Partnering : Technology Modes of Cooperation and Sectoral Differences. *Strategic Management Journal*, 1993 *14*(5), 371–385.

HAGEDOORN, J. Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. *Research Policy*, 2002 *31*(4), 477–492.
HAGEDOORN, J. Sharing intellectual property rights--an exploratory study of joint patenting amongst companies. *Industrial and Corporate Change*, 2003 *12*(5), 1035–1050.

HAGEDOORN, J., & DUYSTERS, G. Learning in Dynamic Inter-Firm Networks: The Efficacy of Multiple Contacts. *Organization Studies*, 2002 *23*(4), 525–548.

HAGEDOORN, J., VAN KRANENBURG, H., & OSBORN, R. N. Joint Patenting Amongst Companies - Exploring the Effects of Inter-Firm R&D Partnering and Experience. *Managerial & Decision Economics*, 2003 *24*(2/3), 71–84.

HAN, Y., & LI, D. Effects of intellectual capital on innovative performance: The role of knowledge based dynamic capability. *Management Decision*, 2015 *53*(1), 40–56.

HARHOFF, D., SCHERER, F. M., & VOPEL, K. (). Citations, family size, opposition and the value of patent rights. *Research Policy*, 2003 *32*(8), 1343–1363.

HARHOFF, D., SCHERER, F. M., & VOPEL, K. Erratum: Citations, family size, opposition and the value of patent rights. *Research Policy*, 2004 *33*(2), 363–364.

HARRISON, S., & SULLIVAN, P. H. Profiting from intellectual capital: Learning from leading companies. *Journal of Intellectual Capital*, 2000 *1*(1), 33–46.

HAUSMAN, J., HALL, B., & GRILICHES, Z. Econometric models for Count Data with application to the patents-R&D relationship. *Econometrica*, 1984 *52*(1), 909–938.

HAUTALA, J. Cognitive proximity in international research groups. *Journal of Knowledge Management*, 2011 *15*(4), 601–624.

HENDERSON, R., & COCKBURN, I. Scale, Scope and Spillovers: The Determinants of Research Productivity in Drug Discovery. *Rand Journal of Economics*, 1996 *27*(1), 32–59.

HILBE, J. . M. *Negative Binomial Regression* (2nd ed.). Cambridge, UK: Cambridge University Press. 2011

HILBE, J. M. *Modeling count data* (1st ed.). New York: Cambridge University Press. 2014

HSU, L. C., & WANG, C. H. Clarifying the Effect of Intellectual Capital on Performance: The Mediating Role of Dynamic Capability. *British Journal of Management*, 2012 *23*(2), 179–205.

HU, A. G. Z., & JAFFE, A. B. (2003). Patent citations and international knowledge flow: The cases of Korea and Taiwan. *International Journal of Industrial Organization*, *21*(6), 849–880.

HU, X., ROSSEAU, R., & CHEN, J. A New Approach for Measuring the Value of Patents Based on Structural Indicators for Ego Patent Citation Networks. *Journal of the American Society for Information Science and Technology*, 2012 *63*(9), 1834–1842.

INKPEN, A. C., & DINUR, A. Knowledge Management Processes and International Joint Ventures. *Organization Science*, 1998 *9*(4), 454–468.

JAFFE, A. B. Technological Oprtunity and Spillovers of R & D : Evidence from Firms ' Patents , Profits and Market Value. *The American Economic Review*, 1986 *76*(5), 984–1001.

JAFFE, A. B., TRAJTENBERG, M., & HENDERSON, R. Geographic Localization of Knowledge. *The Quarterly Journal of Economics*, 1993 *108*(3), 577–598.

JANSEN, J. J. P., TEMPELAAR, M. P., VAN DEN BOSCH, F. A. J., & VOLBERDA, H. W. Structural Differentiation and Ambidexterity: The Mediating Role of Integration Mechanisms. *Organization Science*, 2009 *20*(4), 797–811.

JIANG, R. J., TAO, Q. T., & SANTORO, M. D. Alliance Portfolio Diversity and Firm Performance. *Strategic Management Journal*, 2010 *31*(10), 1136–1144.

JUNNI, P., SARALA, R. M., TARAS, V., & TARBA, S. Y. Organizational ambidexterity and performance: A meta-analysis. *The Academy of Management Perspectives*, 2013 *27*(4), 299–312.

KAISER, U. Measuring knowledge spillovers in manufacturing and services: an empirical assessment of alternative approaches. *Research Policy*, 2002 *31*(1), 125–144.

KALE, P., & SINGH, H Learning and Protection of Proprietary Assets in Strategic Alliances: Building Relational Capital. *Strategic Management Journal*, 2000 *21*(3), 217.

KATILA, R., & AHUJA, G. Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction. *Academy of Management*, *45*(6), 1183–1194. 2002

KIM, C., & INKPEN, A. Cross-border R&D alliances, absorptive capacity and technology learning. *Journal of International Management*, *11*(3), 313–329. 2005

KIM, C., & SONG, J. Creating new technology through alliances: An empirical investigation of joint patents. *Technovation*, *27*(8), 461–470. 2007

KIM, D. H. The link between individual and organizational learning. *Sloan Management Review*, *33*(1), 37–50. 1993

KIM, J. W., & LEE, H. K. Embodied and disembodied international spillovers of R&D in OECD manufacturing industries. *Technovation*, *24*(4), 359–368. 2004

KIM, L. Crisis Construction and Organizational Learning: Capability Building in Catching-up at Hyundai Motor. *Organization Science*, *9*(4), 506–521. 1998

KING, D. R., SLOTEGRAAF, R. J., & KESNER, I. Performance Implications of Firm Resource Interactions in the Acquisition of R&D-Intensive Firms. *OrganizationScience, 19(2), 327–340.* 2008

KNOBEN, J., & OERLEMANS, L. A. G. Proximity and inter-organizational collaboration: A literature review. International Journal of Management Reviews, 8(2), 71–89. 2006a

KNOBEN, J., & OERLEMANS, L. A. G. Proximity and inter-organizational collaboration: A literature review. International Journal of Management Reviews, 8(2), 71–89. 2006b

KOGUT, B. Joint ventures: Theoretical and empirical perspectives. Strategic Management Journal, 9(4), 319–332. 1988

KOGUT, B., & ZANDER, U. Knowledge of the firm, combinative capabilities, and the replication of technology. Organization Science, 3(3), 383–397. 1992

KOGUT, B., & ZANDER, U. What Firms Do? Coordination, Identity, and Learning. Organization Science, 7(5), 502–518. 1996

KYRIAKOPOULOS, K., & MOORMAN, C. Tradeoffs in marketing exploitation and exploration strategies: The overlooked role of market orientation. International Journal of Research in Marketing, 21(3), 219–240. 2004

LANCASTER, G. Research Methods in Managament: A concise introduction to research in management and business consultancy. (G. Lancaster, Ed.) (1st ed., Vol. 1). London, Thousand Oaks, New Delhi: Elsevier Butterworth-Heinemann. 2005

LANE, P. J., KOKA, B. R., & PATHAK, S. The Reification of Absorptive Capacity: A Critical Review and Rejuvenation of the Construct. The Academy of Management Review, 31(4), 833–863. 2006

LANE, P. J., & LUBATKIN, M. Relative Absorptive Capacity and Interorganizational Learning. Strategic Management Journal, 19(5), 461–477. 1998

Lane, P. J., Salk, J. E., & Lyles, M. A. Absorptive capacity, learning, and performance in international joint ventures. Strategic Management Journal, 22(12), 1139–1161. 2001

LANJOUW, J. O., & SCHANKERMAN, M. The Quality of Ideas: Measuring Innovation With Multiple Indicators (No. 7345). NBER working paper. Cambridge, Massachusetts. 1999

LAURSEN, K., LEONE, M. I., & TORRISI, S. Technological exploration through licensing: new insights from the licensee's point of view. Industrial and Corporate Change, 19(3), 871–897. 2010

LAURSEN, K., & SALTER, A. Open for innovation: The role of openness in explaining innovation performance among U.K. manufacturing firms. Strategic Management Journal, 27(2), 131–150. 2006

LAVIE, D., & MILLER, S. R. Alliance portfolio internationalization and firm performance. Organization Science, 19(4), 623–646. 2008

LEE, J., & VELOSO, F. M. Interfirm innovation under uncertainty: Empirical evidence for strategic knowledge partitioning. Journal of Product Innovation Management, 25(5), 418–435. 2008

LEE, W. L., CHIANG, J. C., WU, Y. H., & LIU, C. H. How knowledge exploration distance influences the quality of innovation. Total Quality Management & Business Excellence, 23(9), 1045–1059. 2012

LEE, Y. G. What affects a patent's value? An analysis of variables that affect technological, direct economic, and indirect economic value: An exploratory conceptual approach. Scientometrics, 79(3), 623–633. 2009

LERNER, J. The Importance of Patent Scope: An Empirical Analysis. The RAND Journal of Economics, 25(2), 319–333. 1994

LETEN, B., BELDERBOS, R., & VAN LOOY, B. Technological diversification, coherence, and performance of firms. *Journal of Product Innovation Management*, *24*(6), 567–579. 2007

LI, Y., & VANHAVERBEKE, W. The effects of inter-industry and country difference in supplier relationships on pioneering innovations. *Technovation*, *29*(12), 843–858. 2009

LICHTENTHALER, U. Absorptive Capacity, Environmental Turbulence, and the Complementarity of Organizational Learning Processes. *Academy of Management Journal*, *52*(4), 822–846. 2009

LICHTENTHALER, U., & LICHTENTHALER, E. Technology Transfer across Organizational Boundaries: Absorptive capacity and desorptive capacity. *California Management Review*, *53*(1), 154–170. 2010

LIN, C., WU, Y.-J., CHANG, C., WANG, W., & LEE, C.-Y. The alliance innovation performance of R&D alliances—the absorptive capacity perspective. *Technovation*, *32*(5), 282–292. 2012

LIN, Z., YANG, H., & DEMIRKAN, I. The performance consequences of ambidexterity in strategic alliance formations: Empirical investigation and computational theorizing. *Management Science*, *53*(10), 1645–1658. 2007

LUBATKIN, M., FLORIN, J., & LANE, P. Learning together and apart: A model of reciprocal interfirm learning. *Human Relations*, *54*(10), 1353–1382. 2001

LUBATKIN, M. H., SIMSEK, Z., LING, Y., & VEIGA, J. F. Ambidexterity and Performance in Small-to Medium-Sized Firms: The Pivotal Role of Top Management Team Behavioral Integration. *Journal of Management*, *32*(5), 646–672. 2006

LUO, X., & DENG, L. Do Birds of a Feather Flock Higher? The Effects of Partner Similarity on Innovation in Strategic Alliances in Knowledge-Intensive Industries. *Journal of Management Studies*, *46*(6), 1005–1030. 2009

MACGARVIE, M. The determinants of international knowledge diffusion as measured by patent citations. *Econmics Letters*, *87*(1), 121–126. 2005

MACIONIS, J. J. *Sociology*. (D. Musslewhite, Ed.) (14th ed.). Boston, Massachusetts: Pearson Education Inc. 2012

MAKRI, M., HITT, M. A., & LANE, P. J. Complementary Technologies, Knowledge Relatedness, and Invention Outcomes in High Technology Mergers and Acquisitions. *Strategic Management Journal*, *31*(December 2015), 602–628. 2010

MARABELLI, M., & NEWELL, S.  Knowing, Power and Materiality: A Critical Review and Reconceptualization of Absorptive Capacity. *International Journal of Management Reviews*, *16*(4), 479–499. 2014

MARCH, J. G. Exploration and Exploitation in Organizational Learning. *Organization Science*, *2*(1), 71–87. 1991

MARIANI, M., & ROMANELLI, M. “Stacking” and “picking” inventions: The,patenting behavior of European inventors. *Research Policy*, *36*(8), 1128–1142. 2007

MARTIN, X., & MITCHELL, W. The influence of local search and performance heuristics on new design introduction in a new product market. *Research Policy*, *26*(7-8), 753–771. 1998

MARTINEZ, H., JAIME, A., & CAMACHO, J. Relative absorptive capacity: A research profiling. *Scientometrics*, *92*(3), 657–674. 2012

MATTES, J. Dimensions of Proximity and Knowledge Bases: Innovation between Spatial and Non-spatial Factors. *Regional Studies*, *46*(8), 1085–1099. 2012

MCKELVEY, M., ALM, H., & RICCABONI, M. Does co-location matter for formal knowledge collaboration in the Swedish biotechnology-pharmaceutical sector? *Research Policy*, *32*(3), 483–501. 2003

MCNAMEE, R. C. Can't see the forest for the leaves: Similarity and distance measures for hierarchical taxonomies with a patent classification example. *Research Policy*, *42*(4), 855–873. 2013

MEEUS, M. T. H., OERLEMANS, L. A. G., & HAGE, J. Patterns of interactive learning in a high-tech region. *Organization Studies*, *22*(1), 145–172. 2001

MEIER, M. Knowledge Management in Strategic Alliances: A Review of Empirical Evidence. *International Journal of Management Reviews*, *13*(1), 1–23. 2011

MEISTER, C., & WERKER, C. Physical and organizational proximity in territorial innovation systems: introduction to the special issue. *Journal of Economic Geography*, *4*(1), 1–2. 2004

MEYER, M. What is special about patent citations? differences between scientific and patent citations. *Scientometrics*, *49*(1), 93–123. 2000

MICHEL, J., & BETTELS, B. Patent citation analysis: A close look at the basic input data from patent research reports. *Scientometrics*, *51*(1), 185–201. 2001

MOWERY, D. C., & OXLEY, J. E. (). Inward technology transfer and competitiveness: the role of national innovation systems. *Cambridge Journal of Economics*, *19*(1), 67–93. 1995

MOWERY, D. C., OXLEY, J. E., & SILVERMAN, B. . Strategic Alliances and Interfirm Knowledge Transfer. *Strategic Management Journal*, *17*(1), 77–91. 1996

MOWERY, D. C., OXLEY, J. E., & SILVERMAN, B. S. Technological overlap and interfirm cooperation: implications for the resource-based view of the firm. *Research Policy*, *27*(5), 507–523. 1998

NAMBISAN, S. Industry technical committees, technological distance, and innovation performance. *Research Policy*, *42*(4), 928–940. 2013

NARIN, F., NOMA, E., & PERRY, R. Patents as indicators of corporate technological strength. *Research Policy*, *16*(2-4), 143–155. 1987

NELSON, R. R., & WINTER, S. G. *An evolutionary theory of economic change*. Cambridge, Massachusetts: Harvard University Press. 1982

NEWMAN, W. R., & HANNA, M. D. An empirical exploration of a technology race. *The RAND Journal of Economics*, *28*(2), 228–247. 2006

NOHRIA, N., & GARCIA-PONT, C. Global strategic linkages and industry structure. *Strategic Management Journal*, *12*(Special Issue: Global Strategy), 105–124. 1991

NONAKA, I., & TAKEUCHI, H. *The knowledge-creating company: how Japanese companies create the dynamics of innovation* (1st ed.). New York: Oxford University Press. 1995

NOOTEBOOM, B *Learning and innovation in organizations and economies*. Oxford, UK: Oxford University Press. 2000

NOOTEBOOM, B., VAN HAVERBEKE, W., DUYSTERS, G., GILSING, V., & VAN DEN OORD, A. Optimal cognitive distance and absorptive capacity. *Research Policy*, *36*(7), 1016–1034. 2007

O SEARCOID, M. *Metric Spaces*. *Numerical Methods for Partial Differential Equations* (Vol. 79). London: Springer Berlin Heidelberg. 2007

OECD. *Framework for Biotechnology Statistics*. Paris. OECD Publishing. 2005

OECD. *Innovation in Firms: a Microeconomic Perspective*. París: OECD Publishing. 2009

OECD. *Science, Technology and Industry Scoreboard 2013: Innovation for Growth*. OECD Publishing. 2013

OERLEMANS, L. A. G., & KNOBEN, J. Configurations of knowledge transfer relations: An empirically based taxonomy and its determinants. *Journal of Engineering and Technology Management*, *27*(1-2), 33–51.

OLSSON, O., & FREY, B. S. Entrepreneurship as recombinant 2010growth. *Small Bussines Economics*, *19*(2), 69–80. 2002

ORLANDO, M. J. Measuring spillovers from industrial R&D: on the importance of geographic and technological proximity. *The RAND Journal of Economics*, *35*(4), 777–786. 2004

ORNAGHI, C. Mergers and innovation in big pharma. *International Journal of Industrial Organization*, *27*(1), 70–79. 2009

PANGARKAR, N. Determinants of alliance duration in uncertain environments: The case of the biotechnology sector. *Long Range Planning*, *36*(3), 269–284. 2003

PARK, H., YOON, J., & KIM, K. Identification and evaluation of corporations for merger and acquisition strategies using patent information and text mining. *Scientometrics*, *97*(3), 883–909. 2013

PARK, S. H., & UNGSON, G. R. Interfirm Rivalry and Managerial Complexity: A Conceptual Framework of Alliance Failure. *Organization Science*, *12*(1), 37–53. 2001

PATEL, P., & PAVITT, K. Large Firms in the Production of the World's Technology: An Important Case of "Non-Globalisation." *Journal of International Business Studies*, *22*(1), 1–21. 1991

PERI, G., & URBAN, D. Catching-up to foreign technology? Evidence on the ``Veblen-Gerschenkron'' effect of foreign investments. *Regional Science and Urban Economics*, *36*(1), 72–98. 2006

PETROCELLI, J. V. Hierachial Multiple Regression in Couselling Research: Common Problems and Possible Remedies. *Measurement and Evaluation in Counseling and Development*, *36*(1), 9–14. 2003

PETRUZZELLI, A. M. The impact of technological relatedness, prior ties, and geographical distance on university–industry collaborations: A joint-patent analysis. *Technovation*, *31*(7), 309–319. 2011

PETRUZZELLI, A. M., ALBINO, V., & CARBONARA, N. External knowledge sources and proximity. *Journal of Knowledge Management*, *13*(5), 301–318. 2009

PFEFFER, J., & NOWAK, P. Joint ventures and interorganizational dependence. *Administrative Science Quarterly*, *21*(3), 398–418. 1976

PHENE, A., FLADMOE-LINDQUIST, K., & MARSH, L. Breakthrough innovations in the US biotechnology industry: The effects of technological space and geographic origin. *Strategic Management Journal*, *27*(4), 369–388. 2006

PITSIS, T., & GUDERGAN, S. Know how? Challenges in knowledge and innovation management in alliances: An introduction. *International Journal of Knowledge Management Studies*, *4*(2), 109–113. 2010

PIZER, W. A., & POPP, D. Endogenizing technological change: Matching empirical evidence to modeling needs. *Energy Economics*, *30*(6), 2754–2770. 2008

PODOLNY, J. M., & STUART, T. E. A Role-Based Ecology of Technological Change. *American Journal of Sociology*, *100*(5), 1224–1260. 1995

PORTER, M. E. The Competitive Advantage of Nations. *Harvard Business Review*, *1*(1), 71–91. 1990

POWELL, W. W., KOPUT, K. W., & SMITH-DOERR, L. Interorganizational Collaboration and the Locus of Innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, *41*(1), 116–145. 1996

QUINTANA-GARCIA, C., & BENAVIDES-VELASCO, C. A. Searching for complementary technological knowledge and downstream competences: clustering and cooperation. *International Journal of Technology Management*, *35*(1/2/3/4), 262–283. 2006

QUINTANA-GARCÍA, C., & BENAVIDES-VELAZCO, C. A. Technological relatedness in interfirm cooperation agreements and the generation of innovations. *Cuadernos de Economía Y Direción de Empresa*, *45*, 43–67. 2010

REBOLLEDO, C., & NOLLET, J. Learning from suppliers in the aerospac industry. *International Journal of Production Economics*, *129*(2), 328–337. 2011

REID, D., BUSSIERE, D., & GREENAWAY, K. Alliance formation issues for knowledge-based enterprises. *International Journal of Management Reviews*, *3*(1), 79–100. 2001

REITZIG, M. What determines patent value?: Insights from the semiconductor industry. *Research Policy*, *32*(1), 13–26. 2003

REITZIG, M. Improving patent valuations for management purposes - Validating new indicators by analyzing application rationales. *Research Policy*, *33*(6-7), 939–957. 2004

REITZIG, M., & WAGNER, S. The hidden costs of outsourcing: evidence from patent data. *Strategic Management Journal*, *31*(11), 1183–1201. 2010

RODAN, S., & GALUNIC, D. C. More Than Network Structure : How Knowledge Heterogeneity Influences Managerial Performance and Innovativeness. *Strategic Management Journal*, *25*(6), 541–556. 2004

ROSENKOPF, L., & ALMEIDA, P. Overcoming local search through alliances and mobility. *Management Science*, *49*(6), 751–766. 2003

ROSENKOPF, L., & NERKAR, A. Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management Journal*, *22*(4), 287–306. 2001

ROTHAERMEL, F. T., & DEEDS, D. L. Exploration and exploitation alliances in biotechnology: A system of new product development. *Strategic Management Journal*, *25*(3), 201–221. 2004

ROWLEY, T., BEHRENS, D., & KRACKHARDT, D. Redundant governance structures: an analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic Management Journal*, *21*(3), 369–386. 2000

SAMMARRA, A., & BIGGIERO, L. Heterogeneity and specificity of inter-firm knowledge flows in innovation networks. *Journal of Management Studies*, *45*(4), 800–829. 2008

SAMPSON, R. C. R&D alliances and firm performance: The impact of technological diversity and alliance organization on innovation. *Academy of Management Journal*, *50*(2), 364–386. 2007

SAPIENZA, H. J., PARHANKANGAS, A., & AUTIO, E. Knowledge relatedness and post-spin-off growth. *Journal of Business Venturing*, *19*(6), 809–829. 2004

SCHAMP, E. W., RENTMAISTER, B., & LO, V. Dimensions of proximity in knowledge-based networks: The cases of investment banking and automobile design. *European Planning Studies*, *12*(5), 607–624. 2004

SCHANKERMAN, M., & PAKES, A. Estimates of the Value of Patent Rights in European Countries During the Post-1950 Period. *Economic Journal*, *96*(December), 1052–1076. 1986

SCHERNGELL, T., & BARBER, M. J. Spatial interaction modelling of cross-region R&D collaborations: Empirical evidence from the 5th EU framework programme. *Papers in Regional Science*, *88*(3), 531–546. 2009

SCHERNGELL, T., & BARBER, M. J. Distinct spatial characteristics of industrial and public research collaborations: evidence from the fifth EU Framework Programme. *Annals of Regional Science*, *46*(2), 247–266. 2011

SCHERNGELL, T., & HU, Y. Collaborative Knowledge Production in China: Regional Evidence from a Gravity Model Approach. *Regional Studies*, *45*(6), 755–772. 2011

SCHETTINO, F., STERLACCHINI, A., & VENTURINI, F. Inventive productivity and patent quality: Evidence from Italian inventors. *Journal of Policy Modeling*, *35*(6), 1043–1056. 2013

SCHILDT, H., KEIL, T., & MAULA, M. The temporal effects of relative and firm-level absorptive capacity on interorganizational learning. *Strategic Management Journal*, *33*(10), 1154–1173. 2012

SCHMOCH, U. *Concept of a technology classification for Country comparison*. Final report to the World Intellectual Property Organization. WIPO. 2008

SCHOENMAKERS, W., & DUYSTERS, G. Learning in strategic technology alliances. *Technology Analysis & Strategic Management*, *18*(2), 245–264. 2006

SCHULZE, A., & BROJERDI, G. J. C. The effect of the distance between partners' knowledge components on collaborative innovation. *European Management Review*, *9*(2), 85–98. 2012

SCHUMPETER, J. A. *The theory of economic development: an inquiry into profits, capital, credit, interest, and the business cycle.* London: Oxford Unibersity Press. 1934

SCOTCHMER, S. Standing on the Shoulders of Giants: Cumulative Research and the Patent Law. *Journal of Economic Perspectives*, *5*(1), 29–41. 1991

SHANNON, C. A Mathematical Theory of Communication. *Bell System Technology Journal*, *27*(July), 379–423, 623–656.

SHIN, J., & JALAJAS, D. Technological relatedness, boundary-spanning combination of knowledge and the impact of innovation: Evidence of an inverted-U relationship. *The Journal of High Technology Management Research*, *21*(2), 87–96. 2010

SIMONIN, B. L. Ambiguity and the process of knowledge transfer in strategic alliances. *Strategic Management Journal*, *20*(7), 595–623. 1999

SIMONIN, B. L. An empirical investigation of the process of knowledge transfer in international strategic alliances. *Journal of International Business Studies*, *35*(5), 407–427. 2004

SIMSEK, Z. Organizational Ambidexterity: Towards a Multilevel Understanding. *Journal of Management Studies*, *46*(4), 597–624. 2009

SINGH, J. Distributed R&D, cross-regional knowledge integration and quality of innovative output. *Research Policy*, *37*(1), 77–96. 2008

SNEED, K. A., & JOHNSON, D. K. N. Selling ideas: The determinants of patent value in an auction environment. *R and D Management*, *39*(1), 87–94. 2009

SOH, P., MAHMOOD, I. P., & MITCHELL, W. Dynamic Inducements in R & D Investment : Market Signals and Network Locations. *Academy of Management Journal*, *47*(6), 907–917. 2004

SPENDER, J. C. Limits to Learning from the West: How Western Management Advice May Prove Limited in Eastern Europe. *The International Executive (1986-1998)*, *34*(5), 389–410. 1992

SPENDER, J., & GRANT, R. M. Knowledge of the firm: overview. *Strategic Management Journal*, *17*(S2), 5–9. 1996

STEENSMA, K. H., & LYLES, M. A. Explaining IJV Survival in a Transitional Economy Thorugh Social Exchange and Knowledge Based Perspective. *Academic Management Journal*, *21*(AUGUST), 831–851. 2000

STELLNER, F. Technological distance measures : Theoretical foundation and empirics. In *DRUID Society Conference 2014* (pp. 1–27). Copenhagen. 2014

STELLNER, F. The Impact of Technological Distance on M & A Target Choice and Transaction Value. *Max Planck Institute for Innovation and Competition*. Max Planck Institute. 2015

STUART, T. E., OZDEMIR, S. Z., & DING, W. W. Vertical alliance networks: The case of university-biotechnology-pharmaceutical alliance chains. *Research Policy*, *36*(4), 477–498. 2007

STUART, T. E., & PODOLNY, J. M. Local search and the evolution of technological capabilities. *Strategic Management Journal*, *17*(S1), 21–38. 1996

SZULANSKI, G. Exploring Internal Stickiness: Impediments to the Transfer of Best Practice Within the Firm. *Strategic Management Journal*, *17*(2), 27–43. 1996

TEECE, D. J. Profiting from technological innovation : Implications for integration, collaboration , licensing and public policy. *Research Policy*, *15*(6), 285–305. 1986

TEECE, D. J. Competition, cooperation, and innovation. Organizational arrengements for regimes of rapid technological progress. *Journal of Economic Behavior & Organization*, *18*(1), 1–25. 1992

TEECE, D. J. Capturing value from knowledge assets: The new economy, markets for know-how, and intangible assets. *California Management Review*, *40*(3), 55–80. 1998

TSAI, K. H. Collaborative networks and product innovation performance: Toward a contingency perspective. *Research Policy*, *38*(5), 765–778. 2009

TSAI, W. Knowledge Transfer in Intraorganizational Networks : Effects of Network Position and Absorptive Capacity on Business Unit Innovation and Performance. *Academy of Management*, *44*(5), 996–1004. 2001

TSAI, W., & Ghoshal, S. Social capital and value creation: The role of intrafirm networks. *Academy of Management Journal*, *41*(4), 464–476. 1998

TSENG, F. M., HSIEH, C. H., PENG, Y. N., & CHU, Y. W. Using patent data to analyze trends and the technological strategies of the amorphous silicon thin-film solar cell industry. *Technological Forecasting and Social Change*, *78*(2), 332–345. 2011

Tushman, M. L., & Anderson, P. Technological Discontinuities and Organiza- tional Environments Philip Anderson. *Administrative Science Quarterly*, *31*(3), 439–465. 1986

TZABBAR, D., AHARONSON, B. S., AMBURGEY, T. L., & AL-LAHAM, A. When is the whole bigger than the sum of its parts? Bundling knowledge stocks for innovative success. *Strategic Organization*, *6*(4), 375–406. 2008

UIS, UNAB, & SETIC. *Plan Estratégico de Ciencia, Tecnología e Innovación de Santander*. Bucaramanga. 2013

VAN DE VRANDE, V., VANHAVERBEKE, W., & DUYSTERS, G. Technology In-Sourcing and the Creation of Pioneering Technologies. *Journal of Product Innovation Management*, *28*(6, SI), 974–987. 2011

VAN WIJK, R., JANSEN, J. J. P., & LYLES, M. A. . Inter- and Intra-Organizational Knowledge Transfer: A Meta-Analytic Review and Assessment of its Antecedents and Consequences. *Journal of Management Studies*, *45*(4), 830–853. 2008

VANHABERBEKE, W., DUYSTERS, G., & BEERKENS, B. Technological Capability Building through Networking Strategies within High-Tech Industries. In *Academy of Management Proceeedings* (pp. F1–F6). Academy of Management. 2002

VASUDEVA, G., & ANAND, J. Unpacking Absorptive Capacity : a Study of Knowledge Utilization from Alliance Portfolios. *Academy of Management Journal*, *54*(3), 611–623. 2011

VERDOLINI, E., & GALEOTTI, M. At home and abroad: An empirical analysis of innovation and diffusion in energy technologies. *Journal of Environmental Economics and Management*, *61*(2), 119–134. 2011

VEUGELERS, R., & CASSIMAN, B. R&D cooperation between firms and universities. Some empirical evidence from Belgian manufacturing. *International Journal of Industrial Organization*, *23*(5-6), 355–379. 2005

VOLBERDA, H. W., FOSS, N. J., & LYLES, M. A. Absorbing the Concept of Absorptive Capacity: How to Realize Its Potential in the Organization Field. *Organization Science*, *21*(4), 931–951. 2010

VON HIPPEL, E. *The sources of innovation*. *International Journal of Forecasting* (Vol. 5). New York: Oxford Unibersity Press. 1988

WAGNER, S., & COCKBURN, I. Patents and the survival of Internet-related IPOs. *Research Policy*, *39*(2), 214–228. 2010

WAMPOLD, B. E., & FREUND, R. D. Use of multiple regression in counseling psychology research: a flexible data-analytic strategy. *Journal of Counseling Psychology*, *34*(4), 372–382. 1987

WEBER, B., & WEBER, C. Corporate venture capital as a means of radical innovation: Relational fit, social capital, and knowledge transfer. *Journal of Engineering and Technology Management*, *24*(1-2), 11–35. 2007

WEICK, K. E. *Sensemaking in organizations* (3rd ed.). London, Thousand Oaks, New Delhi: Sage Publications. 1995

WIDENER, S. K. Human capital, pay structure, and the use of performance measures in bonus compensation. *Management Accounting Research*, *17*(2), 198–221. 2006

WIPO. *Intercambiar valor: negociación de acuerdos de licencia de tecnología*. World Intellectual Property Organization. 2005

WIPO. International Patent Classification Guide. World Intellectual Property Organization. 2015

WUYTS, S., COLOMBO, M., DUTTA, S., & NOOTEBOOM, B. Empirical tests of optimal cognitive distance. *Journal of Economic Behavior & Organization*, *58*(2), 277–302. 2005

YLI-RENKO, H., AUTIO, E., & SAPIENZA, H. J. Social capital, knowledge acquisition, and knowledge exploitation in young technology-based firms. *Strategic Management Journal*, *22*(6-7), 587–613. 2001

YOSHINO, M. Y., & RANGAN, S. *Strategic Alliances: An Entrepreneurial Approach to Globalization*. Boston, Massachusetts: Harvard Business School Press. 1995

YOUNDT, M. A., SUBRAMANIAM, M., & SNELL, S. A. Intellectual capital profiles: an examination of investments and returns. *Journal of Management Studies*, *41*(2), 335–361. 2004

ZAHRA, S. A., & GEORGE, G. Absorptive capacity: A Review, Reconceptualization, and Extension. *Academy of Management Review*, *27*(2), 185–203. 2002

ZHAO, Y., LU, Y., & WANG, X. Organizational unlearning and organizational relearning: a dynamic process of knowledge management. *Journal of Knowledge Management*, *17*(6), 902–912. 2013

ZIDORN, W., & WAGNER, M. ). The effect of alliances on innovation patterns: An analysis of the biotechnology industry. *Industrial and Corporate Change*, *22*(6), 1497–1524. 2013

# APPENDICES

## APPENDIX A BIOTECHNOLOGY IPC CODES

| IPC codes | Title |
|---|---|
| A01H 1/00 | Processes for modifying genotypes |
| A01H 4/00 | Plant reproduction by tissue culture techniques |
| A61K 38/00 | Medicinal preparations containing peptides |
| A61K 39/00 | Medicinal preparations containing antigens or antibodies |
| A61K 48/00 | Medicinal preparations containing genetic material which is inserted into cells of the living body to treat genetic diseases; Gene therapy |
| C02F 3/34 | Biological treatment of water, waste water, or sewage: characterized by the micro-organisms used |
| C07G 11/00 | Compounds of unknown constitution: antibiotics |
| C07G 13/00 | Compounds of unknown constitution: vitamins |
| C07G 15/00 | Compounds of unknown constitution: hormones |
| C07K 4/00 | Peptides having up to 20 amino acids in an undefined or only partially defined sequence; Derivatives thereof |
| C07K 14/00 | Peptides having more than 20 amino acids; Gastrins; Somatostatins; Melanotropins; Derivatives thereof |
| C07K 16/00 | Immunoglobulins, e.g. monoclonal or polyclonal antibodies |
| C07K 17/00 | Carrier-bound or immobilized peptides; Preparation thereof |
| C07K 19/00 | Hybrid peptides |
| C12M | Apparatus for enzymology or microbiology |
| C12N | Micro-organisms or enzymes; compositions thereof |
| C12P | Fermentation or enzyme-using processes to synthesize a desired chemical compound or composition or to separate optical isomers from a racemic mixture |
| C12Q | Measuring or testing processes involving enzymes or micro-organisms; compositions or test papers therefor; processes of preparing such compositions; condition-responsive control in microbiological or enzymological processes |
| C12S | Processes using enzymes or micro-organisms to liberate, separate or purify a pre-existing compound or composition processes using enzymes or micro-organisms to treat textiles or to clean solid surfaces of materials |
| G01N 27/327 | Investigating or analyzing materials by the use of electric, electro-chemical, or magnetic means: biochemical electrodes |
| G01N 33/53* | Investigating or analyzing materials by specific methods not covered by the preceding groups: immunoassay; biospecific binding assay; materials therefore |
| G01N 33/54* | Investigating or analyzing materials by specific methods not covered by the preceding groups: double or second antibody: with steric inhibition or signal modification: with an insoluble carrier for immobilizing immunochemicals: the carrier being organic: synthetic resin: as water suspendable particles: with antigen or antibody attached to the carrier via a bridging agent: Carbohydrates: with antigen or antibody entrapped within |

| IPC codes | Title |
|---|---|
| | the carrier |
| G01N 33/55* | Investigating or analyzing materials by specific methods not covered by the preceding groups: the carrier being inorganic: Glass or silica: Metal or metal coated: the carrier being a biological cell or cell fragment: Red blood cell: Fixed or stabilized red blood cell: using kinetic measurement: using diffusion or migration of antigen or antibody: through a gel |
| G01N 33/57* | Investigating or analyzing materials by specific methods not covered by the preceding groups: for venereal disease: for enzymes or isoenzymes: for cancer: for hepatitis: involving monoclonal antibodies: involving limulus lysate |
| G01N 33/68 | Investigating or analyzing materials by specific methods not covered by the preceding groups: involving proteins, peptides or amino acids |
| G01N 33/74 | Investigating or analyzing materials by specific methods not covered by the preceding groups: involving hormones |
| G01N 33/76 | Investigating or analyzing materials by specific methods not covered by the preceding groups: human chorionic gonadotropin |
| G01N 33/78 | Investigating or analyzing materials by specific methods not covered by the preceding groups: thyroid gland hormones |
| G01N 33/88 | Investigating or analyzing materials by specific methods not covered by the preceding groups: involving prostaglandins |
| G01N 33/92 | Investigating or analyzing materials by specific methods not covered by the preceding groups: involving lipids, e.g. cholesterol |

# APPENDIX B. BIOTECHNOLOGY PATENT SAMPLE

| Application Year | Application Number |
|---|---|
| 2006 | US2006011031W 20060327 |
| 2006 | US54294306A 20061004 |
| 2006 | US2006013031W 20060406 |
| 2006 | US2006046625W 20061204 |
| 2006 | US2006015407W 20060424 |
| 2006 | US48424706A 20060710 |
| 2006 | US51753006A 20060906 |
| 2006 | US41054006A 20060425 |
| 2006 | US2006016457W 20060428 |
| 2006 | US2006031277W 20060809 |
| 2006 | US99690206A 20060717 |
| 2006 | US99382306A 20060621 |
| 2006 | US8822806A 20060929 |
| 2006 | US9183506A 20061026 |
| 2006 | US35119006A 20060209 |
| 2006 | US35999706A 20060221 |
| 2006 | US2006060152W 20061023 |
| 2006 | US98839606A 20060630 |
| 2006 | US9306906A 20061110 |
| 2006 | US81427006A 20060119 |
| 2006 | US2006014284W 20060413 |
| 2006 | US2006023607W 20060616 |
| 2006 | US58194506A 20061016 |
| 2006 | US2006019651W 20060518 |
| 2006 | US61175106A 20061215 |
| 2006 | US91330506A 20060526 |
| 2006 | US38624906A 20060321 |
| 2006 | US88460506A 20060216 |
| 2006 | US2006034894W 20060907 |
| 2006 | US6583206A 20060907 |
| 2006 | US99571406A 20060728 |
| 2006 | US34794006A 20060206 |
| 2006 | US35571806A 20060216 |
| 2006 | US35863306A 20060221 |
| 2006 | US35863506A 20060221 |
| 2006 | US36447206A 20060228 |
| 2006 | US36671506A 20060302 |
| 2006 | US43966106A 20060524 |
| 2006 | US91217106A 20060421 |
| 2006 | US2006019414W 20060518 |
| 2006 | US2006042601W 20061101 |
| 2006 | US2006017044W 20060502 |
| 2006 | US2006036268W 20060914 |
| 2006 | US2006034666W 20060907 |
| 2006 | US47000406A 20060905 |
| 2006 | US91758406A 20060616 |
| 2006 | US9787606A 20061227 |
| 2006 | US9787606A 20061227 |
| 2007 | US2007085242W 20071120 |
| 2007 | US28042607A 20070223 |
| 2007 | US98177107A 20071031 |
| 2007 | US2007084276W 20071109 |
| 2007 | US51380907A 20071109 |
| 2007 | US96259407A 20071221 |
| 2007 | US2007011641W 20070514 |
| 2007 | US2007017540W 20070806 |
| 2007 | US37573407A 20070806 |
| 2007 | US94031707A 20071114 |
| 2007 | US81245907A 20070619 |
| 2007 | US94915307A 20071203 |
| 2007 | US2007068982W 20070515 |
| 2007 | US74897807A 20070515 |
| 2007 | US80274207A 20070524 |
| 2007 | US72934107A 20070328 |
| 2007 | US29476207A 20070328 |
| 2007 | US37365907A 20070713 |
| 2007 | US76112607A 20070611 |
| 2007 | US76113007A 20070611 |
| 2007 | US84095207A 20070818 |
| 2007 | US93241007A 20071031 |
| 2007 | US51791907A 20071211 |
| 2007 | US87922607A 20070717 |
| 2007 | US43945907A 20070831 |
| 2007 | US75126107A 20070521 |
| 2007 | US68506807A 20070312 |
| 2007 | US2007003802W 20070209 |
| 2007 | US2007009146W 20070411 |
| 2007 | US2007008597W 20070405 |
| 2007 | US96625807A 20071228 |
| 2007 | US2007081879W 20071019 |
| 2007 | US2007081884W 20071019 |
| 2007 | US92760907A 20071029 |
| 2007 | US85416007A 20070912 |
| 2007 | US88177507A 20070727 |
| 2007 | US78341907A 20070409 |
| 2007 | US92971407A 20071030 |

| Application Year | Application Number |
|---|---|
| 2007 | US69002607A 20070322 |
| 2007 | US77423607A 20070706 |
| 2007 | US2007017748W 20070809 |
| 2007 | US2007025455W 20071213 |
| 2007 | US2007063703W 20070309 |
| 2007 | US90432007A 20070926 |
| 2007 | US2007069347W 20070521 |
| 2007 | US2007061927W 20070209 |
| 2007 | US37566207A 20071001 |
| 2007 | US82348107A 20070627 |
| 2007 | US97866607A 20071030 |
| 2007 | US93303007A 20071031 |
| 2007 | US75663807A 20070601 |
| 2007 | US2007025219W 20071210 |
| 2007 | US75564407A 20070530 |
| 2007 | US2007064106W 20070315 |
| 2007 | US73632207A 20070417 |
| 2007 | US2007025975W 20071220 |
| 2007 | US29561407A 20070330 |
| 2007 | US67935207A 20070227 |
| 2007 | US67946407A 20070227 |
| 2007 | US67948107A 20070227 |
| 2007 | US67950807A 20070227 |
| 2007 | US67952107A 20070227 |
| 2007 | US67960007A 20070227 |
| 2007 | US69114807A 20070326 |
| 2007 | US69135207A 20070326 |
| 2007 | US69137707A 20070326 |
| 2007 | US69227307A 20070328 |
| 2007 | US69538107A 20070402 |
| 2007 | US69538807A 20070402 |
| 2007 | US69548007A 20070402 |
| 2007 | US69551007A 20070402 |
| 2007 | US69551807A 20070402 |
| 2007 | US69750807A 20070406 |
| 2007 | US69752207A 20070406 |
| 2007 | US69752907A 20070406 |
| 2007 | US69755407A 20070406 |
| 2007 | US73519107A 20070413 |
| 2007 | US73806907A 20070420 |
| 2007 | US73808707A 20070420 |
| 2007 | US73810907A 20070420 |

| Application Year | Application Number |
|---|---|
| 2007 | US73813707A 20070420 |
| 2007 | US73815407A 20070420 |
| 2007 | US73820107A 20070420 |
| 2007 | US73984607A 20070425 |
| 2007 | US73992107A 20070425 |
| 2007 | US73994107A 20070425 |
| 2007 | US84307907A 20070822 |
| 2007 | US84309107A 20070822 |
| 2007 | US2007000405W 20070104 |
| 2007 | US2007076160W 20070817 |
| 2007 | US31283707A 20071206 |
| 2007 | US84026707A 20070817 |
| 2007 | US94102207A 20071115 |
| 2007 | US44436107A 20070411 |
| 2007 | US2007082164W 20071022 |
| 2007 | US90082707A 20070913 |
| 2007 | US2007088631W 20071221 |
| 2008 | US28460508A 20080922 |
| 2008 | US67741908A 20080905 |
| 2008 | US10459508A 20080417 |
| 2008 | US52164408A 20080118 |
| 2008 | US9819108A 20080404 |
| 2008 | US66944708A 20080627 |
| 2008 | US2008086808W 20081215 |
| 2008 | US74559708A 20081204 |
| 2008 | US27688908A 20081124 |
| 2008 | US21031308A 20080915 |
| 2008 | US27791908A 20081125 |
| 2008 | US68053908A 20080926 |
| 2008 | US27588508A 20081121 |
| 2008 | US13074008A 20080530 |
| 2008 | US2008008613W 20080715 |
| 2008 | US2008080177W 20081016 |
| 2008 | US10623408A 20080418 |
| 2008 | US12073608A 20080515 |
| 2008 | US34662208A 20081230 |
| 2008 | US67109708A 20080730 |
| 2008 | US74767008A 20081215 |
| 2008 | US2008059125W 20080402 |
| 2008 | US59428908A 20080402 |
| 2008 | US1360608A 20080114 |
| 2008 | US74596208A 20081128 |

| Application Year | Application Number | Application Year | Application Number |
|---|---|---|---|
| 2008 | US2008065659W 20080603 | 2008 | US4661108A 20080312 |
| 2008 | US2008003735W 20080321 | 2008 | US5802208A 20080328 |
| 2008 | US53251708A 20080321 | 2008 | US2008086417W 20081211 |
| 2008 | US2991708A 20080212 | 2008 | US16952708A 20080708 |
| 2008 | US2008059045W 20080401 | 2008 | US2008058486W 20080327 |
| 2008 | US20170508A 20080829 | 2008 | US32982008A 20081208 |
| 2008 | US59984308A 20080516 | 2008 | US11127708A 20080429 |
| 2008 | US68160408A 20080801 | 2008 | US11128408A 20080429 |
| 2008 | US74610908A 20081205 | 2008 | US11133208A 20080429 |
| 2008 | US14850708A 20080418 | 2008 | US11149608A 20080429 |
| 2008 | US2905408A 20080211 | 2008 | US11151808A 20080429 |
| 2008 | US2008004016W 20080327 | 2008 | US11154008A 20080429 |
| 2008 | US2008079904W 20081015 | 2008 | US11158508A 20080429 |
| 2008 | US27023508A 20081113 | 2008 | US11218408A 20080430 |
| 2008 | US59701408A 20080510 | 2008 | US11221208A 20080430 |
| 2008 | US14827508A 20080417 | 2008 | US11222308A 20080430 |
| 2008 | US9979808A 20080409 | 2008 | US13734508A 20080611 |
| 2008 | US14226008A 20080619 | 2008 | US4027908A 20080229 |
| 2008 | US59867008A 20080509 | 2008 | US4129708A 20080303 |
| 2008 | US1010708A 20080118 | 2008 | US4216008A 20080304 |
| 2008 | US33507108A 20081215 | 2008 | US5456408A 20080325 |
| 2008 | US67249808A 20080806 | 2008 | US5793908A 20080328 |
| 2008 | US67575808A 20080620 | 2008 | US5886808A 20080331 |
| 2008 | US2008011089W 20080925 | 2008 | US6144608A 20080402 |
| 2008 | US28488808A 20080925 | 2008 | US6213508A 20080403 |
| 2008 | US11447708A 20080502 | 2008 | US60154408A 20080509 |
| 2008 | US11175008A 20080429 | 2008 | US24522308A 20081003 |
| 2008 | US13264208A 20080604 | 2008 | US2008057694W 20080320 |
| 2008 | US12357208A 20080520 | 2008 | US53142608A 20080320 |
| 2008 | US33333408A 20081212 | 2008 | US74666608A 20081206 |
| 2008 | US33333508A 20081212 | 2008 | US2008003751W 20080321 |
| 2008 | US33345508A 20081212 | 2008 | US4653908A 20080312 |
| 2008 | US33348408A 20081212 | 2008 | US2008060486W 20080416 |
| 2008 | US19700508A 20080822 | 2008 | US74565208A 20081202 |
| 2008 | US37566808A 20080122 | 2008 | US33294408A 20081211 |
| 2008 | US53294108A 20080326 | 2008 | US2008001721W 20080208 |
| 2008 | US10538008A 20080418 | 2008 | US2008012821W 20081114 |
| 2008 | US25256108A 20081016 | 2008 | US6928408A 20080208 |
| 2008 | US33085108A 20081209 | 2008 | US2008080531W 20081020 |
| 2008 | US4518308A 20080310 | 2009 | US62808509A 20091130 |
| 2008 | US60212708A 20080505 | 2009 | US2009005082W 20090910 |
| 2008 | US59892208A 20080502 | 2009 | US48387109A 20090612 |
| 2008 | US52608508A 20080208 | 2009 | US2009038158W 20090325 |

| Application Year | Application Number |
|---|---|
| 2009 | US2009038164W 20090325 |
| 2009 | US50723709A 20090722 |
| 2009 | US32196309A 20090126 |
| 2009 | US57652209A 20091009 |
| 2009 | US2009043908W 20090514 |
| 2009 | US46482909A 20090512 |
| 2009 | US73629509A 20090327 |
| 2009 | US58268909A 20091020 |
| 2009 | US200913060154A 20090820 |
| 2009 | US2009044376W 20090518 |
| 2009 | US58518009A 20090908 |
| 2009 | US2009043667W 20090512 |
| 2009 | US99573309A 20090602 |
| 2009 | US200913318660A 20091203 |
| 2009 | US200913147007A 20090130 |
| 2009 | US63251709A 20091207 |
| 2009 | US2009066146W 20091130 |
| 2009 | US2009042640W 20090502 |
| 2009 | US92084409A 20090309 |
| 2009 | US2009063457W 20091105 |
| 2009 | US48966309A 20090623 |
| 2009 | US57849909A 20091013 |
| 2009 | US2009032425W 20090129 |
| 2009 | US99544209A 20090601 |
| 2009 | US200913122239A 20091007 |
| 2009 | US200913003499A 20090707 |
| 2009 | US2009049429W 20090701 |
| 2009 | US2009057021W 20090915 |
| 2009 | US2009057037W 20090915 |
| 2009 | US2009069862W 20091230 |
| 2009 | US49657309A 20090701 |
| 2009 | US56031709A 20090915 |
| 2009 | US56039009A 20090915 |
| 2009 | US2009066392W 20091202 |
| 2009 | US50448709A 20090716 |
| 2009 | US2009044399W 20090518 |
| 2009 | US46780109A 20090518 |
| 2009 | US46780809A 20090518 |
| 2009 | US46782009A 20090518 |
| 2009 | US46782609A 20090518 |
| 2009 | US2009050985W 20090717 |
| 2009 | US50516609A 20090717 |
| 2009 | US41768309A 20090403 |

| Application Year | Application Number |
|---|---|
| 2009 | US42952909A 20090424 |
| 2009 | US42587409A 20090417 |
| 2009 | US54527909A 20090821 |
| 2009 | US200913000288A 20090715 |
| 2009 | US35692309A 20090121 |
| 2009 | US2009003897W 20090630 |
| 2009 | US38914009A 20090219 |
| 2009 | US57380109A 20091005 |
| 2009 | US45770809A 20090603 |
| 2009 | US92112909A 20090227 |
| 2009 | US200913266773A 20091203 |
| 2009 | US54823909A 20090826 |
| 2009 | US60679209A 20091027 |
| 2009 | US40202909A 20090311 |
| 2009 | US46322009A 20090508 |
| 2009 | US46790509A 20090518 |
| 2009 | US53511909A 20090804 |
| 2009 | US53515209A 20090804 |
| 2009 | US2009033042W 20090204 |
| 2009 | US60460009A 20091023 |
| 2009 | US93549909A 20090331 |
| 2009 | US42942809A 20090424 |
| 2009 | US43097509A 20090428 |
| 2009 | US43099609A 20090428 |
| 2009 | US43101309A 20090428 |
| 2009 | US43110709A 20090428 |
| 2009 | US43441109A 20090501 |
| 2009 | US43443109A 20090501 |
| 2009 | US43446709A 20090501 |
| 2009 | US43449409A 20090501 |
| 2009 | US43523309A 20090504 |
| 2009 | US43708409A 20090507 |
| 2009 | US46304809A 20090508 |
| 2009 | US46310309A 20090508 |
| 2009 | US47122209A 20090522 |
| 2009 | US47122809A 20090522 |
| 2009 | US53533809A 20090804 |
| 2009 | US53536609A 20090804 |
| 2009 | US53625909A 20090805 |
| 2009 | US53628609A 20090805 |
| 2009 | US53632909A 20090805 |
| 2009 | US53635709A 20090805 |
| 2009 | US54096709A 20090813 |

| Application Year | Application Number | Application Year | Application Number |
|---|---|---|---|
| 2009 | US54131209A 20090814 | 2010 | US80575210A 20100818 |
| 2009 | US54687709A 20090825 | 2010 | US78956410A 20100528 |
| 2009 | US54689709A 20090825 | 2010 | US71927510A 20100308 |
| 2009 | US54730509A 20090825 | 2010 | US2010037487W 20100604 |
| 2009 | US86442409A 20090126 | 2010 | US201013517109A 201002 |
| 2009 | US99441109A 20090519 | 2010 | US201013394069A 201009 |
| 2009 | US99920109A 20090519 | 2010 | US2010032134W 20100422 |
| 2009 | US200913138223A 20090122 | 2010 | US2010038904W 20100616 |
| 2009 | US32211109A 20090129 | 2010 | US81709310A 20100616 |
| 2009 | US2009069868W 20091230 | 2010 | US81713410A 20100616 |
| 2009 | US200913054527A 20090717 | 2010 | US97657210A 20101222 |
| 2009 | US2009065355W 20091120 | 2010 | US97302010A 20101220 |
| 2009 | US2009046837W 20090610 | 2010 | US2010057862W 20101123 |
| 2009 | US48099109A 20090609 | 2010 | US95309810A 20101123 |
| 2009 | US200913001389A 20090701 | 2010 | US94850310A 20101117 |
| 2009 | US2009049895W 20090708 | 2010 | US2010054930W 20101101 |
| 2009 | US46211909A 20090728 | 2010 | US72655410A 20100318 |
| 2009 | US2009050807W 20090716 | 2010 | US94555110A 20101112 |
| 2009 | US2009034511W 20090219 | 2010 | US2010044522W 20100805 |
| 2010 | US2010030314W 20100408 | 2010 | US2010045174W 20100811 |
| 2010 | US2010026429W 20100305 | 2010 | US85133010A 20100805 |
| 2010 | US71890210A 20100305 | 2010 | US70974810A 20100222 |
| 2010 | US86294810A 20100825 | 2010 | US2010048532W 20100910 |
| 2010 | US70568310A 20100215 | 2010 | US2010047132W 20100830 |
| 2010 | US91039310A 20101022 | 2010 | US87134510A 20100830 |
| 2010 | US95009410A 20101119 | 2010 | US201013503220A 201010 |
| 2010 | US75641710A 20100408 | 2010 | US201013509968A 201011 |
| 2010 | US2010051960W 20101008 | 2010 | US2010043182W 20100726 |
| 2010 | US70596210A 20100215 | 2010 | US201013388028A 201007 |
| 2010 | US85924610A 20100818 | 2010 | US96237510A 20101207 |
| 2010 | US2010054553W 20101028 | 2010 | US201013147153A 201001 |
| 2010 | US201013260702A 20100326 | 2010 | US91461510A 20101028 |
| 2010 | US78539610A 20100521 | 2010 | US201013319054A 201005 |
| 2010 | US70752710A 20100217 | 2010 | US201013258737A 201004 |
| 2010 | US94414510A 20101111 | 2010 | US75753210A 20100409 |
| 2010 | US201013509210A 20101112 | 2010 | US75756810A 20100409 |
| 2010 | US2010057869W 20101123 | 2010 | US75757610A 20100409 |
| 2010 | US2010057886W 20101123 | 2010 | US75758810A 20100409 |
| 2010 | US201013511629A 20101123 | 2010 | US76012510A 20100414 |
| 2010 | US95321210A 20101123 | 2010 | US76015810A 20100414 |
| 2010 | US95325110A 20101123 | 2010 | US76017910A 20100414 |
| 2010 | US69220810A 20100122 | 2010 | US76023010A 20100414 |

| Application Year | Application Number |
|---|---|
| 2010 | US76026410A 20100414 |
| 2010 | US76204210A 20100416 |
| 2010 | US76205510A 20100416 |
| 2010 | US76206710A 20100416 |
| 2010 | US76474210A 20100421 |
| 2010 | US76475710A 20100421 |
| 2010 | US76603410A 20100423 |
| 2010 | US76604410A 20100423 |
| 2010 | US76608110A 20100423 |
| 2010 | US76608810A 20100423 |
| 2010 | US76800410A 20100427 |
| 2010 | US76801510A 20100427 |
| 2010 | US76803010A 20100427 |
| 2010 | US77011610A 20100429 |
| 2010 | US77016210A 20100429 |
| 2010 | US77156410A 20100430 |
| 2010 | US77160610A 20100430 |
| 2010 | US77582910A 20100507 |
| 2010 | US77585810A 20100507 |
| 2010 | US77587910A 20100507 |
| 2010 | US77594810A 20100507 |
| 2010 | US77770710A 20100511 |
| 2010 | US77774110A 20100511 |
| 2010 | US77778010A 20100511 |
| 2010 | US77780210A 20100511 |
| 2010 | US77786110A 20100511 |
| 2010 | US78042810A 20100514 |
| 2010 | US201013518298A 20101217 |
| 2010 | US2010059578W 20101208 |
| 2010 | US95174410A 20101122 |
| 2010 | US77630310A 20100507 |
| 2010 | US71204210A 20100224 |
| 2010 | US80252710A 20100607 |
| 2010 | US2010044206W 20100803 |
| 2010 | US71349710A 20100226 |

# APPENDIX C RESEARCH R® MARKDOWN ANALYSES

# Research R Markdown

Hugo Martínez

October 2015

This is a Markdown document created from the statistical language **R**. It includes both content as well as the output of the main embedded **R** code used in the research within the document.

```
setwd("~/R/Doctorado/Versión 3")
```

---

## 1. READING DATA

1. Importing the data to **R** platform

```
doct<-read.table("Muestra Final_V_0.1.txt", header=T, na.string="N.A")
```

---

## 2. DATA SCREENING

Basic statistics

### 2.1 Dependent variable

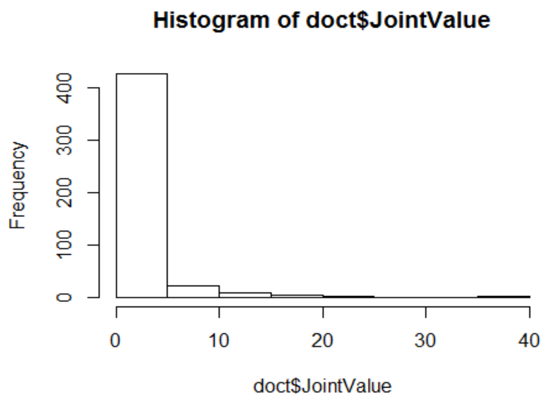
```
library(pastecs)
library(ggplot2)
library(GGally)
library(reshape)
library(lme4)
library(compiler)
library(parallel)
library(boot)

stat.desc(doct$JointValue)

##      nbr.val     nbr.null       nbr.na          min          max
##  465.0000000  268.0000000    0.0000000    0.0000000   38.0000000
##        range          sum       median         mean      SE.mean
##   38.0000000  777.0000000    0.0000000    1.6709677    0.1607975
## CI.mean.0.95          var      std.dev     coef.var
##    0.3159816   12.0229700    3.4674155    2.0750942
```

The mean is 1.6709 and the variance is 12.022 Var>>mean.

```
table(doct$JointValue)

##
##   0   1   2   3   4   5   6   7   8   9  10  11  12  16  17  20  21  38
## 268  64  35  25  17  18   5   6   5   6   1   3   6   1   2   1   1   1
```

It is important to have in mind the number of zeros=268. Is a count data variable¡¡¡

```r
hist(doct$JointValue)
```




Can be seen the frequency distribution of the dependent variable. A high percentage of zeros is found: 57.63441%.

### 2.1 Logarithm of geographical distance

```r
#GeoDist= Distancia geográfica entre firmas,

doct$GeoDistance<-log(doct$GeoDistance)
```

### 2.2 Creating the General Descriptive Statistics table

```r
#library(stargazer)
#stargazer(doct[c("JointValue","TeachDistCos",  "StuDistCos",  "TeachCap",  "Stu
Cap",  "GeoDistance",  "PriorCollCount",  "NumCod","PatCitT","PCT" )], type="tex
t", title="Descriptive Statistics",align=T, digits=3)
```

The generated table of descriptive statistics is included in the document.

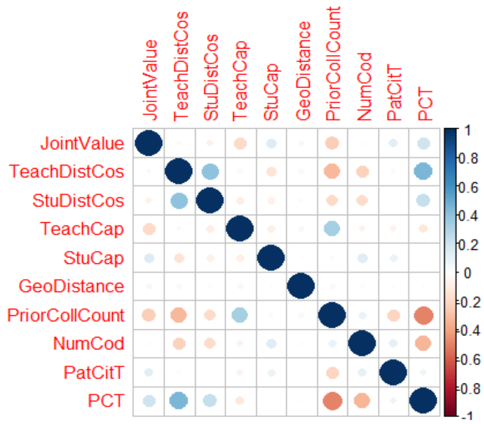
### 2.3 Creating the Correlation graphic representation

```r
correlation<-cor(doct[ ,c("JointValue","TeachDistCos", "StuDistCos", "TeachCa
p", "StuCap", "GeoDistance", "PriorCollCount", "NumCod","PatCitT","PCT")])


#stargazer(correlation,    type="text", title="Correlation Table", align=TRUE,
digits=3)
```

The generated table of correlation between variables is included in the document.

It is worth to create a graphic correlation figure or table in order to understand better the data

```r
#First graphic representation of correlation
library(corrplot)

corrplot(cor(doct[,c("JointValue","TeachDistCos", "StuDistCos", "TeachCap", "
StuCap", "GeoDistance", "PriorCollCount", "NumCod","PatCitT","PCT")]))
```




### 2.1 Data trends and Outliers

Now, to check "Outliers"

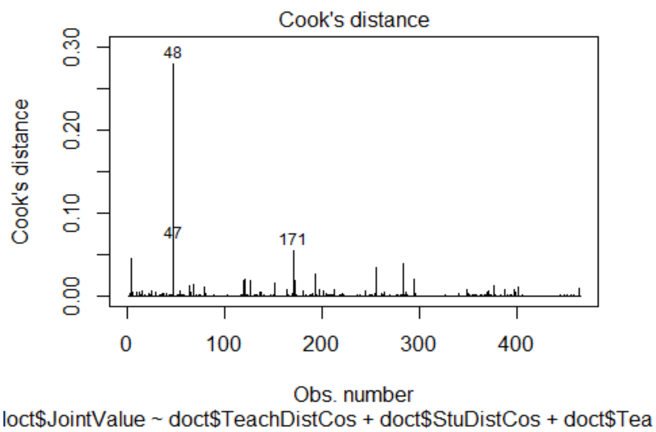
```
library(car)
library(COUNT)
fit<-glm(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap+doct
$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$NumCo
d+doct$X2007+doct$X2008+doct$X2009+doct$X2010, data=doct)
outlierTest(fit) # p-valor de Bonferonni para las observaciones mas extremas

##      rstudent unadjusted p-value Bonferonni p
## 48  12.176514         4.1463e-34   1.9280e-31
## 171  6.180193         6.4023e-10   2.9771e-07
## 47   4.986673         6.1428e-07   2.8564e-04
## 255  4.499005         6.8272e-06   3.1747e-03
## 284  4.209892         2.5549e-05   1.1880e-02
```

After the outlier test, some observations seems to be outliers. However, it has not assessed yet whether it influences the regression. Though, after checking the observations was not possible to find any reason to exclude them from the sample. To be sure about their behavior, a further analysis was made.

Then, now a check for Influential Observations using the Cook Distance and Influence Plot was made.

```
cutoff <- 4/((nrow(doct)-length(fit$coefficients)-2)) # Identificamos los val
ores D > 4/(n-k-1)
cutoff

## [1] 0.008908686
```

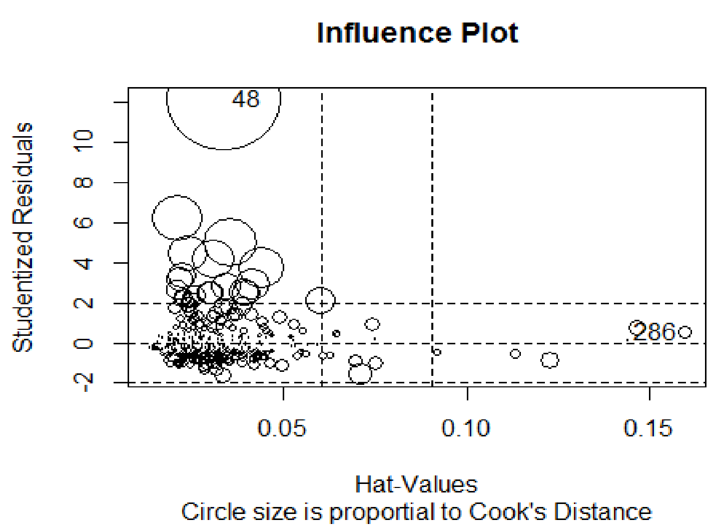
```r
plot(fit, which=4, cook.levels=cutoff) #Gráfica de la distancia de Cook
```




According to the results, the cut-off line is 0.009. Some observations have an unusually high level of influence in the regression.

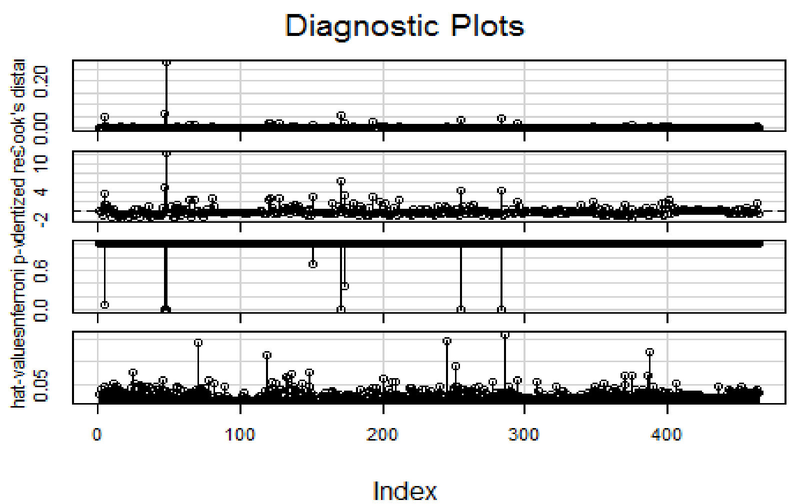
In order to display the studentized residuals, hat-values and Cook´s distance it is helpful to analyze an Influence Plot.

```r
influencePlot(fit,  id.method="Aleatorio", main="Influence Plot", sub="Circle size is proportial to Cook's Distance" ) #Gráfica de observaciones que influencian.
```

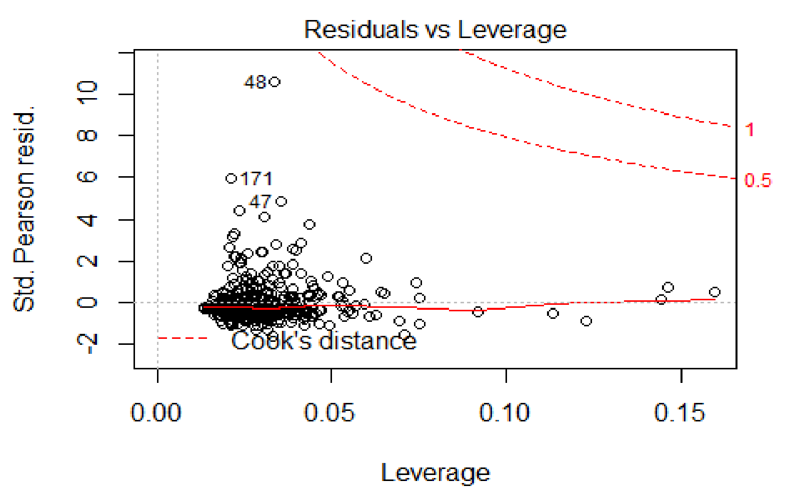



```
##         StudRes        Hat      CookD
## 48   12.1765142 0.03376957 0.52822867
## 286   0.5399634 0.15957080 0.06293146
```

```
influenceIndexPlot(fit) #Muestra Indexplots, levarage, distancia de Cook, res
iduos studiantized, niveles de significancia para Outliers.
```




```
plot(fit, which=5)
```




Although it seems some apparent outlier points, once checked I didn't find any strong reason why I have to remove them out from the model. Then, I decided to run the model with the actual data including a model with robust regression estimators which weight outlying data.

### 2.5 Variance Inflation Factors (VIFs) test

Now, having into account the correlation matrix, it is important to check for multicollinearity.

```
library(car)
fit<-glm(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap+doct
$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$NumCo
d+doct$X2007+doct$X2008+doct$X2009+doct$X2010, data=doct)
Testvif<-vif(fit)
Testvif
```

```
##    doct$TeachDistCos      doct$StuDistCos        doct$TeachCap
##             1.541341             1.236550             1.210376
##          doct$StuCap     doct$GeoDistance doct$PriorCollCount
##             1.062130             1.038768             1.882102
##         doct$PatCitT             doct$PCT         doct$NumCod
##             1.123881             1.748787             1.210932
##          doct$X2007           doct$X2008           doct$X2009
##             2.442897             2.514761             2.626376
##          doct$X2010
##             2.671201
```

```
Testvif<10   #Observamos si el VIF de cada variable es menor de 10
```

```
##    doct$TeachDistCos      doct$StuDistCos        doct$TeachCap
##                 TRUE                 TRUE                 TRUE
##          doct$StuCap     doct$GeoDistance doct$PriorCollCount
##                 TRUE                 TRUE                 TRUE
##         doct$PatCitT             doct$PCT         doct$NumCod
##                 TRUE                 TRUE                 TRUE
##          doct$X2007           doct$X2008           doct$X2009
##                 TRUE                 TRUE                 TRUE
##          doct$X2010
##                 TRUE
```

```
max(Testvif)   #Encontramos el máximo VIF
```

```
## [1] 2.671201
```

```
mean(Testvif)   #Encontramos el promedio de VIF
```

```
## [1] 1.716162
```

```
  sqrt(vif(fit)) > 2 # problema?
```

```
##    doct$TeachDistCos      doct$StuDistCos       doct$TeachCap
##                FALSE                FALSE               FALSE
##          doct$StuCap     doct$GeoDistance doct$PriorCollCount
##                FALSE                FALSE               FALSE
##         doct$PatCitT             doct$PCT         doct$NumCod
##                FALSE                FALSE               FALSE
##          doct$X2007           doct$X2008          doct$X2009
##                FALSE                FALSE               FALSE
##          doct$X2010
##                FALSE
```

According to the test there is not a VIF>10, the max VIF is 2.6712011¡¡¡. Therefore, I do not expect multicollinearity problems.

To be completely sure about the potential Multicollinearity problems, I check for Condition Index of the variables:

```
library(perturb)
  colldiag(fit)

## Condition
## Index    Variance Decomposition Proportions
##             intercept doct$TeachDistCos doct$StuDistCos doct$TeachCap
## 1     1.000 0.001     0.004             0.005           0.004
## 2     2.205 0.000     0.017             0.018           0.006
## 3     2.517 0.000     0.002             0.004           0.000
## 4     2.536 0.000     0.000             0.002           0.003
## 5     2.687 0.000     0.000             0.000           0.008
## 6     2.974 0.000     0.017             0.031           0.009
## 7     3.670 0.000     0.000             0.000           0.085
## 8     3.992 0.000     0.020             0.239           0.020
## 9     4.315 0.000     0.016             0.453           0.046
## 10    5.353 0.002     0.051             0.007           0.469
## 11    5.669 0.000     0.494             0.139           0.278
## 12    6.087 0.000     0.292             0.069           0.012
## 13    6.804 0.018     0.065             0.006           0.005
## 14   14.438 0.978     0.022             0.028           0.056
##    doct$StuCap doct$GeoDistance doct$PriorCollCount doct$PatCitT doct$PCT
## 1  0.006       0.003            0.002               0.004        0.003
## 2  0.001       0.000            0.092               0.001        0.032
## 3  0.001       0.000            0.024               0.053        0.001
## 4  0.002       0.001            0.000               0.071        0.001
## 5  0.001       0.000            0.043               0.136        0.021
## 6  0.033       0.000            0.003               0.231        0.031
## 7  0.165       0.000            0.060               0.412        0.006
## 8  0.323       0.000            0.003               0.000        0.073
## 9  0.263       0.008            0.029               0.023        0.079
## 10 0.073       0.349            0.053               0.001        0.016
## 11 0.007       0.119            0.276               0.039        0.000
## 12 0.045       0.185            0.240               0.009        0.614
## 13 0.036       0.117            0.170               0.017        0.075
## 14 0.044       0.215            0.005               0.003        0.047
##    doct$NumCod doct$X2007 doct$X2008 doct$X2009 doct$X2010
## 1  0.005       0.001      0.002      0.001      0.001
## 2  0.016       0.002      0.026      0.000      0.034
## 3  0.009       0.085      0.003      0.140      0.002
## 4  0.000       0.125      0.017      0.019      0.093
## 5  0.001       0.005      0.157      0.023      0.016
## 6  0.095       0.011      0.023      0.030      0.053
## 7  0.188       0.004      0.002      0.001      0.023
## 8  0.238       0.002      0.008      0.001      0.000
## 9  0.131       0.005      0.002      0.004      0.002
## 10 0.052       0.012      0.014      0.002      0.001
## 11 0.001       0.045      0.006      0.000      0.001
## 12 0.111       0.019      0.015      0.000      0.017
## 13 0.091       0.301      0.343      0.358      0.368
## 14 0.061       0.383      0.382      0.419      0.390
```

From the results, the Max Condition Index is 14.438 which is lower than the threshold of 30, therefore there are not problems with Multicollinearity.

## 2. Poisson Model Analysis

### 2.1 Regression, dispersion parameter

```
library(COUNT)
#setwd("~/R/Doctorado")

disp<-glm(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap+doc
t$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$NumC
od+doct$X2007+doct$X2008+doct$X2009+doct$X2010 , family=poisson)
summary(disp)

##
## Call:
## glm(formula = doct$JointValue ~ doct$TeachDistCos + doct$StuDistCos +
##     doct$TeachCap + doct$StuCap + doct$GeoDistance + doct$PriorCollCount +

##     doct$PatCitT + doct$PCT + doct$NumCod + doct$X2007 + doct$X2008 +
##     doct$X2009 + doct$X2010, family = poisson)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.6974  -1.5395  -0.8032   0.0119   9.4716
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          1.2148887  0.1617964   7.509 5.97e-14 ***
## doct$TeachDistCos   -0.7683646  0.1380222  -5.567 2.59e-08 ***
## doct$StuDistCos     -0.6366512  0.1278018  -4.982 6.31e-07 ***
## doct$TeachCap       -0.0047729  0.0009600  -4.972 6.64e-07 ***
## doct$StuCap          0.0045791  0.0008637   5.302 1.15e-07 ***
## doct$GeoDistance     0.0323972  0.0126743   2.556   0.0106 *
## doct$PriorCollCount -0.0270628  0.0033475  -8.084 6.24e-16 ***
## doct$PatCitT         0.0043089  0.0009682   4.450 8.57e-06 ***
## doct$PCT             0.4678622  0.0921776   5.076 3.86e-07 ***
## doct$NumCod         -0.0081480  0.0073626  -1.107   0.2684
## doct$X2007          -0.4526510  0.1131759  -4.000 6.35e-05 ***
## doct$X2008          -0.2384904  0.1088672  -2.191   0.0285 *
## doct$X2009          -0.4600448  0.1139010  -4.039 5.37e-05 ***
## doct$X2010          -1.0087874  0.1359835  -7.418 1.18e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1984.5  on 464  degrees of freedom
## Residual deviance: 1389.7  on 451  degrees of freedom
## AIC: 1987.4
##
## Number of Fisher Scoring iterations: 6
```

At first glance, most of the variables are significant except for the variable NumCod. However, the model has to be checked further:

```
#Calculamos el estadístico de dispersión
pr <- sum(residuals(disp, type="pearson")^2)   # Pearson Chi2
pr/disp$df.residual    #Dispersion   statistic

## [1] 4.199037
```

The dispersion statistic is more than 1. The model seems to be overdisperse.

```
#Algunos estadísticos del ajuste del modelo
modelfit(disp)

## $AIC
## [1] 1987.445
##
## $AICn
## [1] 4.274075
##
## $BIC
## [1] 2045.433
##
## $BICqh
## [1] 4.372771
```

Estadísticos para comparar modelos.

### 2.1 Frequency table and number of zeros

```
#Tabla de frecuencias, porcentajes, y acumulados
cnt <- table(doct$JointValue)
dataf <- data.frame(prop.table(table(doct$JointValue) ) )
dataf$cumulative <- cumsum(dataf$Freq)
datafall <- data.frame(cnt, dataf$Freq*100, dataf$cumulative * 100)
datafall

##    Var1 Freq dataf.Freq...100 dataf.cumulative...100
## 1     0  268       57.6344086               57.63441
## 2     1   64       13.7634409               71.39785
## 3     2   35        7.5268817               78.92473
## 4     3   25        5.3763441               84.30108
## 5     4   17        3.6559140               87.95699
## 6     5   18        3.8709677               91.82796
## 7     6    5        1.0752688               92.90323
## 8     7    6        1.2903226               94.19355
## 9     8    5        1.0752688               95.26882
## 10    9    6        1.2903226               96.55914
## 11   10    1        0.2150538               96.77419
## 12   11    3        0.6451613               97.41935
## 13   12    6        1.2903226               98.70968
## 14   16    1        0.2150538               98.92473
## 15   17    2        0.4301075               99.35484
## 16   20    1        0.2150538               99.56989
## 17   21    1        0.2150538               99.78495
## 18   38    1        0.2150538              100.00000
```

It is evident again, the high percentage of zeros: 57.63%. It would be helpful to know what is the expected number of zeros in a Poisson model for the observed mean.

```
#número de ceros esperado
mu<-mean(doct$JointValue)
Numzeros<-exp(-mu)
Numzeros

## [1] 0.188065
```

It is expected that a probability of 18.8% of the observations in the model have a zero count with a mean of 1.67. However, 57.63% of zero counts in the data. There is a big difference between the observed against the expected.

### 2.1 Predicted versus Observed Counts and variance comparison

```
#Observamos la diferencia en porcentajes de las proporciones con la media pre
dicha y la media observada
poi.obs.pred(38, model=disp)

##     Count   propObsv     propPred          Diff
## 1       0 57.6344086 3.635762e+01  2.127679e+01
## 2       1 13.7634409 2.217197e+01 -8.408534e+00
## 3       2  7.5268817 1.576592e+01 -8.239035e+00
## 4       3  5.3763441 1.036948e+01 -4.993132e+00
## 5       4  3.6559140 6.399383e+00 -2.743469e+00
## 6       5  3.8709677 3.800965e+00  7.000245e-02
## 7       6  1.0752688 2.207815e+00 -1.132546e+00
## 8       7  1.2903226 1.263503e+00  2.682008e-02
## 9       8  1.0752688 7.151849e-01  3.600839e-01
## 10      9  1.2903226 4.022719e-01  8.880506e-01
## 11     10  0.2150538 2.266292e-01 -1.157542e-02
## 12     11  0.6451613 1.292786e-01  5.158826e-01
## 13     12  1.2903226 7.543629e-02  1.214886e+00
## 14     13  0.0000000 4.520644e-02 -4.520644e-02
## 15     14  0.0000000 2.766776e-02 -2.766776e-02
## 16     15  0.0000000 1.706867e-02 -1.706867e-02
## 17     16  0.2150538 1.045112e-02  2.046026e-01
## 18     17  0.4301075 6.266774e-03  4.238408e-01
## 19     18  0.0000000 3.644842e-03 -3.644842e-03
## 20     19  0.0000000 2.043996e-03 -2.043996e-03
## 21     20  0.2150538 1.101664e-03  2.139521e-01
## 22     21  0.2150538 5.698757e-04  2.144839e-01
## 23     22  0.0000000 2.828530e-04 -2.828530e-04
## 24     23  0.0000000 1.347634e-04 -1.347634e-04
## 25     24  0.0000000 6.168243e-05 -6.168243e-05
## 26     25  0.0000000 2.714973e-05 -2.714973e-05
## 27     26  0.0000000 1.150436e-05 -1.150436e-05
## 28     27  0.0000000 4.698354e-06 -4.698354e-06
## 29     28  0.0000000 1.851426e-06 -1.851426e-06
## 30     29  0.0000000 7.047323e-07 -7.047323e-07
## 31     30  0.0000000 2.593957e-07 -2.593957e-07
## 32     31  0.0000000 9.242015e-08 -9.242015e-08
## 33     32  0.0000000 3.190510e-08 -3.190510e-08
## 34     33  0.0000000 1.068186e-08 -1.068186e-08
## 35     34  0.0000000 3.471455e-09 -3.471455e-09
## 36     35  0.0000000 1.096020e-09 -1.096020e-09
## 37     36  0.0000000 3.364453e-10 -3.364453e-10
## 38     37  0.0000000 1.004913e-10 -1.004913e-10
## 39     38  0.2150538 2.922630e-11  2.150538e-01
```

```r
#Observamos la diferencia en las varianzas
disp<-glm(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap+doc
t$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$NumC
od+doct$X2007+doct$X2008+doct$X2009+doct$X2010, family=poisson)
summary(disp)

##
## Call:
## glm(formula = doct$JointValue ~ doct$TeachDistCos + doct$StuDistCos +
##     doct$TeachCap + doct$StuCap + doct$GeoDistance + doct$PriorCollCount +
##     doct$PatCitT + doct$PCT + doct$NumCod + doct$X2007 + doct$X2008 +
##     doct$X2009 + doct$X2010, family = poisson)
##


## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.6974  -1.5395  -0.8032   0.0119   9.4716
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          1.2148887  0.1617964   7.509 5.97e-14 ***
## doct$TeachDistCos   -0.7683646  0.1380222  -5.567 2.59e-08 ***
## doct$StuDistCos     -0.6366512  0.1278018  -4.982 6.31e-07 ***
## doct$TeachCap       -0.0047729  0.0009600  -4.972 6.64e-07 ***
## doct$StuCap          0.0045791  0.0008637   5.302 1.15e-07 ***
## doct$GeoDistance     0.0323972  0.0126743   2.556   0.0106 *
## doct$PriorCollCount -0.0270628  0.0033475  -8.084 6.24e-16 ***
## doct$PatCitT         0.0043089  0.0009682   4.450 8.57e-06 ***
## doct$PCT             0.4678622  0.0921776   5.076 3.86e-07 ***
## doct$NumCod         -0.0081480  0.0073626  -1.107   0.2684
## doct$X2007          -0.4526510  0.1131759  -4.000 6.35e-05 ***
## doct$X2008          -0.2384904  0.1088672  -2.191   0.0285 *
## doct$X2009          -0.4600448  0.1139010  -4.039 5.37e-05 ***
## doct$X2010          -1.0087874  0.1359835  -7.418 1.18e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1984.5  on 464  degrees of freedom
## Residual deviance: 1389.7  on 451  degrees of freedom
## AIC: 1987.4
##
## Number of Fisher Scoring iterations: 6

xbp<-predict(disp)
mup<-exp(xbp)
mean(mup) #Varianza esperada

## [1] 1.670968

var(doct$JointValue)   #Varianza observada

## [1] 12.02297
```

For the Poisson model, the observed variance is 12.02297, while the expected variance is 1.6709677. Seems to have Overdispersion!!!

### 2.1 Comparing between model Standard errors, Scaled Standard Errors, Robust Standard Errors

1. Scaled Standard Errors

```
#Errores Estandar del Modelo
disp<-glm(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap+doc
t$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$NumC
od+doct$X2007+doct$X2008+doct$X2009+doct$X2010, family=poisson)
summary(disp)
```

```
##
## Call:
## glm(formula = doct$JointValue ~ doct$TeachDistCos + doct$StuDistCos +
##     doct$TeachCap + doct$StuCap + doct$GeoDistance + doct$PriorCollCount +

##     doct$PatCitT + doct$PCT + doct$NumCod + doct$X2007 + doct$X2008 +
##     doct$X2009 + doct$X2010, family = poisson)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.6974  -1.5395  -0.8032   0.0119   9.4716
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)          1.2148887  0.1617964   7.509 5.97e-14 ***
## doct$TeachDistCos   -0.7683646  0.1380222  -5.567 2.59e-08 ***
## doct$StuDistCos     -0.6366512  0.1278018  -4.982 6.31e-07 ***
## doct$TeachCap       -0.0047729  0.0009600  -4.972 6.64e-07 ***
## doct$StuCap          0.0045791  0.0008637   5.302 1.15e-07 ***
## doct$GeoDistance     0.0323972  0.0126743   2.556   0.0106 *
## doct$PriorCollCount -0.0270628  0.0033475  -8.084 6.24e-16 ***
## doct$PatCitT         0.0043089  0.0009682   4.450 8.57e-06 ***
## doct$PCT             0.4678622  0.0921776   5.076 3.86e-07 ***
## doct$NumCod         -0.0081480  0.0073626  -1.107   0.2684
## doct$X2007          -0.4526510  0.1131759  -4.000 6.35e-05 ***
## doct$X2008          -0.2384904  0.1088672  -2.191   0.0285 *
## doct$X2009          -0.4600448  0.1139010  -4.039 5.37e-05 ***
## doct$X2010          -1.0087874  0.1359835  -7.418 1.18e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 1984.5  on 464  degrees of freedom
## Residual deviance: 1389.7  on 451  degrees of freedom
## AIC: 1987.4
##
## Number of Fisher Scoring iterations: 6
```

```
confint(disp)
```

```
## Waiting for profiling to be done...
```

```
##                              2.5 %       97.5 %
## (Intercept)           0.894151062   1.528505436
## doct$TeachDistCos    -1.041197797 -0.500007864
## doct$StuDistCos      -0.889857795 -0.388655033
## doct$TeachCap        -0.006689667 -0.002925124
## doct$StuCap           0.002875090   0.006261870
## doct$GeoDistance      0.008000354   0.057715238
## doct$PriorCollCount -0.034038955 -0.020883271
## doct$PatCitT          0.002350484   0.006149950
## doct$PCT              0.288949029   0.650414834
## doct$NumCod          -0.022951834   0.005933088
## doct$X2007           -0.674728117 -0.230668219
## doct$X2008           -0.451364042 -0.024259336
## doct$X2009           -0.683318245 -0.236444631
## doct$X2010           -1.278512478 -0.744835257
```

```
pr <- sum(residuals(disp,type="pearson")^2)       # Estadístico de Pearson
dispersion <- pr/disp$df.residual; dispersion           # dispersión
```

```
## [1] 4.199037
```

```
#Errores Estándar Escalados del Modelo
sse <- sqrt(diag(vcov(disp))) * sqrt(dispersion); sse
```

```
##         (Intercept)    doct$TeachDistCos      doct$StuDistCos
##         0.331545822          0.282828800          0.261885740
##       doct$TeachCap          doct$StuCap     doct$GeoDistance
##         0.001967229          0.001769868          0.025971683
## doct$PriorCollCount         doct$PatCitT             doct$PCT
##         0.006859574          0.001983974          0.188886285
##          doct$NumCod           doct$X2007           doct$X2008
##         0.015087067          0.231914965          0.223085864
##          doct$X2009           doct$X2010
##         0.233400772          0.278651300
```

```
#Coeficientes del modelo con errores escalados
dispQL<-glm(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap+d
oct$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$Nu
mCod+doct$X2007+doct$X2008+doct$X2009+doct$X2010, family=quasipoisson)
coef(dispQL); confint(dispQL)
```

```
##         (Intercept)    doct$TeachDistCos      doct$StuDistCos
##         1.214888683         -0.768364585         -0.636651168
##       doct$TeachCap          doct$StuCap     doct$GeoDistance
##        -0.004772864          0.004579091          0.032397226
## doct$PriorCollCount         doct$PatCitT             doct$PCT
##        -0.027062809          0.004308931          0.467862224
##          doct$NumCod           doct$X2007           doct$X2008
##        -0.008148038         -0.452651003         -0.238490415
##          doct$X2009           doct$X2010
##        -0.460044797         -1.008787396
```

```
## Waiting for profiling to be done...
```

```
##                              2.5 %        97.5 %
## (Intercept)            0.549582947   1.8502903554
## doct$TeachDistCos     -1.332753410 -0.2227720827
## doct$StuDistCos       -1.161886019 -0.1333449580
## doct$TeachCap         -0.008778990 -0.0010560712
## doct$StuCap            0.001060841   0.0080081783
## doct$GeoDistance      -0.016711381   0.0853741907
## doct$PriorCollCount -0.042314420 -0.0151396477
## doct$PatCitT           0.000155987   0.0079692524
## doct$PCT               0.104712677   0.8463029966
## doct$NumCod           -0.039337057   0.0200094854
## doct$X2007            -0.909197519   0.0034975465
## doct$X2008            -0.674391503   0.2031295068
## doct$X2009            -0.918498928 -0.0002253077
## doct$X2010            -1.569828993 -0.4720949720
```

There is a considerable difference between the Standard Errors of model and Standard Errors adjusted by escalating them by the dispersion parameter. Scaled standard errors are about the double-triple of standard errors.

1. Robust Standard Errors

```
#Calculo de errores robustos o empiricos
library(sandwich)
disp<-glm(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap+doc
t$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$NumC
od+doct$X2007+doct$X2008+doct$X2009+doct$X2010, family=poisson)
# Matriz de Varianza Covarianza
sqrt(diag(vcovHC(disp, type="HC0")))                  # final HC0 = H-C-zero
```

```
##          (Intercept)   doct$TeachDistCos     doct$StuDistCos
##          0.387220877         0.331803357         0.281921078
##        doct$TeachCap          doct$StuCap    doct$GeoDistance
##          0.002138888         0.002056822         0.025383752
## doct$PriorCollCount         doct$PatCitT            doct$PCT
##          0.005413015         0.001790218         0.233519736
##          doct$NumCod          doct$X2007          doct$X2008
##          0.015590772         0.296641344         0.299956573
##           doct$X2009          doct$X2010
##          0.292407968         0.323154919
```

Here, also is found a difference between standard errors and robust standard errors. Robust standard errors are about the double-triple of standard errors. This can be taken as a proof of the high overdispersion of the Poisson model

### 4. Negative Binomial Regression

Now a calculation of the different models for the inclusion of the different variables.

#### 4.1 Inclusion of regression Variables

##### 4.1.1 Model 1: Base Model: Control Variables

```
library(gamlss)
#Cálculo de las variables de control
disp1<-glm.nb(doct$JointValue~doct$GeoDistance+doct$PriorCollCount+doct$PatCi
tT+doct$PCT+doct$NumCod+doct$X2007+doct$X2008+doct$X2009+doct$X2010,  data=doc
t)
summary(disp1)
```

```
##
## Call:
## glm.nb(formula = doct$JointValue ~ doct$GeoDistance + doct$PriorCollCount
+
##     doct$PatCitT + doct$PCT + doct$NumCod + doct$X2007 + doct$X2008 +
##     doct$X2009 + doct$X2010, data = doct, init.theta = 0.4468866606,
##     link = log)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.43660  -1.14987  -0.46753  -0.05127   2.67308
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           0.501511   0.360015   1.393   0.1636
## doct$GeoDistance      0.033609   0.028548   1.177   0.2391
## doct$PriorCollCount -0.028327   0.004457  -6.355 2.08e-10 ***
## doct$PatCitT          0.005838   0.002648   2.204   0.0275 *
## doct$PCT              0.419380   0.205885   2.037   0.0417 *
## doct$NumCod           0.015504   0.016503   0.939   0.3475
## doct$X2007           -0.405595   0.298730  -1.358   0.1745
## doct$X2008           -0.253129   0.289384  -0.875   0.3817
## doct$X2009           -0.311251   0.298651  -1.042   0.2973
## doct$X2010           -0.772925   0.321315  -2.406   0.0161 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.4469) family taken to be 1)
##
##     Null deviance: 507.92  on 464  degrees of freedom
## Residual deviance: 380.96  on 455  degrees of freedom
## AIC: 1415.7
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.4469
##           Std. Err.:  0.0506
##
##  2 x log-likelihood:  -1393.6660
```

#### 4.1.1 Model 2: Including TeachDistCos

```
disp2<-glm.nb(doct$JointValue~doct$TeachDistCos+doct$GeoDistance+doct$PriorCo
llCount+doct$PatCitT+doct$PCT+doct$NumCod+doct$X2007+doct$X2008+doct$X2009+do
ct$X2010, data=doct)
summary(disp2)
```

```
##
## Call:
## glm.nb(formula = doct$JointValue ~ doct$TeachDistCos + doct$GeoDistance +
##     doct$PriorCollCount + doct$PatCitT + doct$PCT + doct$NumCod +
##     doct$X2007 + doct$X2008 + doct$X2009 + doct$X2010, data = doct,
##     init.theta = 0.4837332286, link = log)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.54950  -1.09074  -0.52752  -0.04129   2.70842
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)           0.710570   0.367928   1.931  0.05345 .
## doct$TeachDistCos    -1.195431   0.304101  -3.931 8.46e-05 ***
## doct$GeoDistance      0.040626   0.028143   1.444  0.14886
## doct$PriorCollCount -0.030222   0.004525  -6.679 2.40e-11 ***
## doct$PatCitT          0.005882   0.002572   2.287  0.02218 *
## doct$PCT              0.608822   0.209491   2.906  0.00366 **
## doct$NumCod           0.007770   0.016281   0.477  0.63321
## doct$X2007           -0.317558   0.293300  -1.083  0.27894
## doct$X2008           -0.030734   0.286245  -0.107  0.91450
## doct$X2009           -0.121065   0.293192  -0.413  0.67966
## doct$X2010           -0.709303   0.315555  -2.248  0.02459 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.4837) family taken to be 1)
##
##     Null deviance: 531.97  on 464  degrees of freedom
## Residual deviance: 383.22  on 454  degrees of freedom
## AIC: 1402.9
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.4837
##           Std. Err.:  0.0563
##
##  2 x log-likelihood:  -1378.9050
```

### 4.1.1 Model 3: Including StuDistCos

```
disp3<-glm.nb(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$GeoDista
nce+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$NumCod+doct$X2007+doct$X20
08+doct$X2009+doct$X2010, data=doct)
summary(disp3)
```

```
##
## Call:
## glm.nb(formula = doct$JointValue ~ doct$TeachDistCos + doct$StuDistCos +
##     doct$GeoDistance + doct$PriorCollCount + doct$PatCitT + doct$PCT +
##     doct$NumCod + doct$X2007 + doct$X2008 + doct$X2009 + doct$X2010,
##     data = doct, init.theta = 0.4998846644, link = log)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.59915  -1.09522  -0.55736  -0.04159   2.46300
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)          9.543e-01  3.719e-01   2.566 0.010297 *
## doct$TeachDistCos   -1.034e+00  3.133e-01  -3.301 0.000964 ***
## doct$StuDistCos     -6.851e-01  2.753e-01  -2.488 0.012844 *
## doct$GeoDistance     3.499e-02  2.788e-02   1.255 0.209416
## doct$PriorCollCount -3.144e-02  4.576e-03  -6.870 6.41e-12 ***
## doct$PatCitT         5.979e-03  2.539e-03   2.355 0.018509 *
## doct$PCT             5.854e-01  2.072e-01   2.825 0.004730 **
## doct$NumCod          4.103e-05  1.640e-02   0.003 0.998003
## doct$X2007          -3.563e-01  2.914e-01  -1.223 0.221337
## doct$X2008          -1.371e-02  2.843e-01  -0.048 0.961552
## doct$X2009          -1.042e-01  2.905e-01  -0.359 0.719731
## doct$X2010          -7.066e-01  3.128e-01  -2.259 0.023901 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.4999) family taken to be 1)
##
##     Null deviance: 542.15  on 464  degrees of freedom
## Residual deviance: 383.75  on 453  degrees of freedom
## AIC: 1398.5
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.4999
##           Std. Err.:  0.0587
##
##  2 x log-likelihood:  -1372.5060
```

### 4.1.1 Model 4: Including TeachCap

```r
disp4<-glm.nb(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap
+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$NumCod+doct$
X2007+doct$X2008+doct$X2009+doct$X2010, data=doct)
summary(disp4)
```

```
##
## Call:
## glm.nb(formula = doct$JointValue ~ doct$TeachDistCos + doct$StuDistCos   +
##     doct$TeachCap + doct$GeoDistance + doct$PriorCollCount +
##     doct$PatCitT + doct$PCT + doct$NumCod  + doct$X2007 + doct$X2008 +
##     doct$X2009 + doct$X2010,  data = doct, init.theta = 0.5156256639,
##     link = log)
##
## Deviance Residuals:
##      Min         1Q    Median        3Q       Max
## -1.65965  -1.07098  -0.55706   0.01815   2.35023
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          1.127285   0.375517   3.002  0.00268 **
## doct$TeachDistCos   -0.979889   0.312390  -3.137  0.00171 **
## doct$StuDistCos     -0.703201   0.272383  -2.582  0.00983 **
## doct$TeachCap       -0.005511   0.001942  -2.838  0.00455 **
## doct$GeoDistance     0.040857   0.027839   1.468  0.14221
## doct$PriorCollCount -0.028192   0.004691  -6.010 1.85e-09 ***
## doct$PatCitT         0.006365   0.002512   2.534  0.01127 *
## doct$PCT             0.608812   0.205880   2.957  0.00311 **
## doct$NumCod         -0.005023   0.016336  -0.307  0.75849
## doct$X2007          -0.320647   0.291710  -1.099  0.27168
## doct$X2008           0.018573   0.283481   0.066  0.94776
## doct$X2009          -0.102511   0.289807  -0.354  0.72355
## doct$X2010          -0.652487   0.310710  -2.100  0.03573 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for  Negative Binomial(0.5156) family taken to be  1)
##
##     Null deviance: 551.87  on 464  degrees of freedom
## Residual deviance: 381.98  on 452  degrees of freedom
## AIC: 1392.3
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.5156
##           Std. Err.:  0.0608
##
##  2 x log-likelihood:  -1364.2730
```

#### 4.1.1 Model 5: Including StuCap: Final Model

```
disp5<-glm.nb(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap
+doct$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$
NumCod+doct$X2007+doct$X2008+doct$X2009+doct$X2010, data=doct)
summary(disp5)
```

```
##
## Call:
## glm.nb(formula = doct$JointValue ~ doct$TeachDistCos + doct$StuDistCos +
##     doct$TeachCap + doct$StuCap + doct$GeoDistance + doct$PriorCollCount +

##     doct$PatCitT + doct$PCT + doct$NumCod + doct$X2007 + doct$X2008 +
##     doct$X2009 + doct$X2010, data = doct, init.theta = 0.5381602379,
##     link = log)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.73290  -1.04278  -0.53834   0.00241   2.52086
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          0.927139   0.379796   2.441  0.01464 *
## doct$TeachDistCos   -0.978738   0.312511  -3.132  0.00174 **
## doct$StuDistCos     -0.602993   0.270099  -2.232  0.02558 *
## doct$TeachCap       -0.004656   0.001936  -2.405  0.01618 *
## doct$StuCap          0.006422   0.002163   2.969  0.00298 **
## doct$GeoDistance     0.046368   0.027648   1.677  0.09353 .
## doct$PriorCollCount -0.029328   0.004790  -6.123 9.16e-10 ***
## doct$PatCitT         0.006319   0.002480   2.548  0.01084 *
## doct$PCT             0.497897   0.204085   2.440  0.01470 *
## doct$NumCod         -0.009265   0.016118  -0.575  0.56543
## doct$X2007          -0.468040   0.286699  -1.633  0.10257
## doct$X2008          -0.051787   0.277385  -0.187  0.85190
## doct$X2009          -0.246299   0.284906  -0.864  0.38732
## doct$X2010          -0.716103   0.306125  -2.339  0.01932 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.5382) family taken to be 1)
##
##     Null deviance: 565.47  on 464  degrees of freedom
## Residual deviance: 382.87  on 451  degrees of freedom
## AIC: 1386.3
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.5382
##           Std. Err.:  0.0643
##
##  2 x log-likelihood:  -1356.3380
```

### 4.1.5 Model 6: Final Model using robust estimators

```
#The use of robust standar errors was made using Stata software.

#library(stargazer)
#stargazer(disp1,disp2,disp3,disp4,disp5,disp6, type="txt", title="Determinan
ts of Innvention Value", no.space=T,digits=4)
```

The table that includes the six regression models is included in the document.

### 4.2 Wald Test to models

Now the Wald test is calculated in order to know the overall effect of the variables in the model.

```
library(aod)
```

#### 4.2.1 Model 1: Base Model: Control Variables

```
wald.test(b=coef(disp1), Sigma=vcov(disp1), Terms=1:10)

## Wald test:
## ----------
##
## Chi-squared test:
## X2 = 124.6, df = 10, P(> X2) = 0.0
```

The chi-squared test statistic of 124.6, with ten degrees of freedom is associated with a p-value of 0.0 indicating that the overall effect of Control Variables is statistically significant.

#### 4.2.2 Model 2: Including TeachDistCos

```
wald.test(b=coef(disp2), Sigma=vcov(disp2), Terms=1:11)

## Wald test:
## ----------
##
## Chi-squared test:
## X2 = 136.5, df = 11, P(> X2) = 0.0
```

The chi-squared test statistic of 136.5, with eleven degrees of freedom is associated with a p-value of 0.0 indicating that the overall effect of including the teacher technological distance is statistically significant.

#### 4.2.3 Model 3: Including StuDistCos

```
wald.test(b=coef(disp3), Sigma=vcov(disp3), Terms=1:12)

## Wald test:
## ----------
##
## Chi-squared test:
## X2 = 142.9, df = 12, P(> X2) = 0.0
```

The chi-squared test statistic of 142.9, with twelve degrees of freedom is associated with a p-value of 0.0 indicating that the overall effect of including the teacher technological distance is statistically significant.

### 4.2.2 Model 4: Including TeachCap

```
wald.test(b=coef(disp4), Sigma=vcov(disp4), Terms=1:13)

## Wald test:
## ----------
##
## Chi-squared test:
## X2 = 157.1, df = 13, P(> X2) = 0.0
```

The chi-squared test statistic of 157.1, with thirteen degrees of freedom is associated with a p-value of 0.0 indicating that the overall effect of including the teacher technological distance is statistically significant.

### 4.2.4 Model 5: Final Model: Including StuCap

```
wald.test(b=coef(disp5), Sigma=vcov(disp5), Terms=1:14)

## Wald test:
## ----------
##
## Chi-squared test:
## X2 = 165.2, df = 14, P(> X2) = 0.0
```

The chi-squared test statistic of 165.2, with fourteen degrees of freedom is associated with a p-value of 0.0 indicating that the overall effect of including the teacher technological distance is statistically significant.

---

## 5 Model with other classes of technological distances

### 5.1 Model 7: Distance based on the Cosine of Similarity

```
disp7<-glm.nb(doct$JointValue~doct$TeachDistCos+doct$StuDistCos+doct$TeachCap
+doct$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$
NumCod+doct$X2007+doct$X2008+doct$X2009+doct$X2010, data=doct)
summary(disp7)

##
## Call:
## glm.nb(formula = doct$JointValue ~ doct$TeachDistCos + doct$StuDistCos +
##     doct$TeachCap + doct$StuCap + doct$GeoDistance + doct$PriorCollCount +

##     doct$PatCitT + doct$PCT + doct$NumCod  + doct$X2007 + doct$X2008 +
##     doct$X2009 + doct$X2010,  data = doct, init.theta = 0.5381602379,
##     link = log)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.73290  -1.04278  -0.53834   0.00241   2.52086
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)           0.927139    0.379796    2.441   0.01464 *
## doct$TeachDistCos    -0.978738    0.312511   -3.132   0.00174 **
## doct$StuDistCos      -0.602993    0.270099   -2.232   0.02558 *
## doct$TeachCap        -0.004656    0.001936   -2.405   0.01618 *
## doct$StuCap           0.006422    0.002163    2.969   0.00298 **
## doct$GeoDistance      0.046368    0.027648    1.677   0.09353 .
## doct$PriorCollCount -0.029328    0.004790   -6.123 9.16e-10 ***
## doct$PatCitT          0.006319    0.002480    2.548   0.01084 *
## doct$PCT              0.497897    0.204085    2.440   0.01470 *
## doct$NumCod          -0.009265    0.016118   -0.575   0.56543
## doct$X2007           -0.468040    0.286699   -1.633   0.10257
## doct$X2008           -0.051787    0.277385   -0.187   0.85190
## doct$X2009           -0.246299    0.284906   -0.864   0.38732
## doct$X2010           -0.716103    0.306125   -2.339   0.01932 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.5382) family taken to be 1)
##
##     Null deviance: 565.47  on 464  degrees of freedom
## Residual deviance: 382.87  on 451  degrees of freedom
## AIC: 1386.3
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.5382
##           Std. Err.:  0.0643
##
##  2 x log-likelihood:  -1356.3380
```

### 5.1 Model 8: Distance based on the Min-Complement measure

```
disp8<-glm.nb(doct$JointValue~doct$TeachDistMincomp+doct$StuDistMincomp+doct$
TeachCap+doct$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$P
CT+doct$NumCod+doct$X2007+doct$X2008+doct$X2009+doct$X2010,    data=doct)
summary(disp8)

##
## Call:
## glm.nb(formula = doct$JointValue ~ doct$TeachDistMincomp + doct$StuDistMin
comp +
##     doct$TeachCap + doct$StuCap + doct$GeoDistance + doct$PriorCollCount +

##     doct$PatCitT + doct$PCT + doct$NumCod  + doct$X2007 + doct$X2008 +
##     doct$X2009 + doct$X2010, data = doct, init.theta = 0.52536795,
##     link = log)
##
## Deviance  Residuals:
##      Min        1Q    Median        3Q       Max
## -1.73414   -1.07835  -0.52383   -0.01886   2.67940
##
## Coefficients:
##                        Estimate  Std. Error  z value  Pr(>|z|)
## (Intercept)            1.183848    0.410628    2.883   0.00394 **
## doct$TeachDistMincomp  -0.857020    0.389271   -2.202   0.02769 *
## doct$StuDistMincomp    -0.710768    0.334188   -2.127   0.03343 *
## doct$TeachCap          -0.004104    0.001934   -2.121   0.03389 *
```

```
## doct$StuCap             0.006300   0.002193   2.873  0.00406 **
## doct$GeoDistance        0.048744   0.027736   1.757  0.07885 .
## doct$PriorCollCount    -0.028748   0.004705  -6.110 9.94e-10 ***
## doct$PatCitT            0.006173   0.002502   2.468  0.01360 *
## doct$PCT                0.418354   0.204193   2.049  0.04048 *
## doct$NumCod            -0.014504   0.016576  -0.875  0.38156
## doct$X2007             -0.466719   0.289130  -1.614  0.10648
## doct$X2008             -0.058733   0.279441  -0.210  0.83353
## doct$X2009             -0.215678   0.287803  -0.749  0.45362
## doct$X2010             -0.601917   0.306737  -1.962  0.04973 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.5254) family taken to be 1)
##
##     Null deviance: 557.8  on 464  degrees of freedom
## Residual deviance: 383.4  on 451  degrees of freedom
## AIC: 1391.8
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.5254
##           Std. Err.:  0.0625
##
##  2 x log-likelihood:  -1361.8130
```

### 5.1 Model 9: Distance based on the Euclidean measure

```
disp9<-glm.nb(doct$JointValue~doct$TeachDistEuc+doct$StuDistEuc+doct$TeachCap
+doct$StuCap+doct$GeoDistance+doct$PriorCollCount+doct$PatCitT+doct$PCT+doct$
NumCod+doct$X2007+doct$X2008+doct$X2009+doct$X2010, data=doct)
summary(disp9)
```

```
##
## Call:
## glm.nb(formula = doct$JointValue ~ doct$TeachDistEuc + doct$StuDistEuc +
##     doct$TeachCap + doct$StuCap + doct$GeoDistance + doct$PriorCollCount +

##     doct$PatCitT + doct$PCT + doct$NumCod + doct$X2007 + doct$X2008 +
##     doct$X2009 + doct$X2010, data = doct, init.theta = 0.5210554108,
##     link = log)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.73074  -1.06600  -0.49176  -0.01608   2.67149
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          1.237578   0.425233   2.910  0.00361 **
## doct$TeachDistEuc   -0.717319   0.317015  -2.263  0.02365 *
## doct$StuDistEuc     -0.676248   0.383025  -1.766  0.07747 .
## doct$TeachCap       -0.004462   0.001922  -2.322  0.02023 *
## doct$StuCap          0.006519   0.002184   2.985  0.00284 **
## doct$GeoDistance     0.046083   0.027740   1.661  0.09667 .
## doct$PriorCollCount -0.028770   0.004694  -6.130 8.81e-10 ***
## doct$PatCitT         0.006251   0.002507   2.494  0.01265 *
## doct$PCT             0.389212   0.204924   1.899  0.05753 .
```

```
## doct$NumCod          -0.015387    0.016783   -0.917   0.35923
## doct$X2007           -0.507290    0.289635   -1.751   0.07986 .
## doct$X2008           -0.116163    0.279260   -0.416   0.67743
## doct$X2009           -0.273328    0.288012   -0.949   0.34261
## doct$X2010           -0.628518    0.307057   -2.047   0.04067 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.5211) family taken to be 1)
##
##     Null deviance: 555.18  on 464  degrees of freedom
## Residual deviance: 383.69  on 451  degrees of freedom
## AIC: 1393.8
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.5211
##           Std. Err.:  0.0619
##
##  2 x log-likelihood:  -1363.8040
```