ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

"A LONGITUDINAL STUDY OF THE FACTORS THAT AFFECT FIRMS' ABILITY TO SHAPE TECHNOLOGICAL EVOLUTION AND TO EXERT A BROAD INFLUENCE ON TECHNOLOGICAL CHANGE"

By:

Michalis E. Papazoglou

A Dissertation Submitted to the Faculty of Athens University of Economics and Business in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Supervided by Associate Professor Yiannis E. Spanos

.

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Department of Management Studies and Technology

December 2015

To my wife Katerina and to my children Lenia and Louiza

Table of Contents

List of Figures and Tables:

CHAPTER 1

INTRODUCTION

1 INTRODUCTION

1.1 Objectives of the Study

Among organization scholars there is a growing interest in the study of the organization-environment co-evolution. The majority of them focus on how organizations are shaped by their environment and only a few examine how organizations systematically influence their environment (Lewin and Volberda 1999). Thinking about this research area, Nelson and Winter (1982), in one of the most influential works in evolutionary social science, posited that “Perhaps in the future it will become possible to build and comprehend models of industry evolution that are based on detailed and realistic models of individual firm behavior” (p. 36). Building on this line of reasoning, the overarching objective of this thesis is to shed light on how specific firm’s characteristics can enhance or limit the ability of a firm to affect technological evolution and to exert a broad influence on technological change.

Generally, the interactions between organizations and their environments has always been a fundamental issue within organization science. In this study, we adopted a broadly co-evolutionary framework where there is a two-way relationship between the organization and its environment. Specifically, it is a framework that takes into account and link together both "downward processes" (i.e. the environment’s impact on organizations) and "upward processes" (i.e. the organizations’ ability to change their strategies and routines in response or in anticipation to changes in their environments –or even pro-active strategies to change the prevailing “rules of the game”-, and the consequential effects on these environments). Under this framework, important research streams in organization studies have emerged, such as the adaptation-selection (Lewin and Volberda 1999) and the agency-structure debate (Heugens and Lander 2009).

However, probably because of the difficulties in detecting and collecting data for documenting the effect of an organization’s activities on its surrounding context (Barley and Tolbert 1997), the extant literature so far has primarily focused on how and why organizations tend to become isomorphic with their environments, whereas questions of how organizations systematically influence their environments are studied less (Baum and Singh 1994b). Although there exist some theoretical approaches that acknowledge, and are either fully based on the shaping role of organizations (e.g. institutional entrepreneurship) or partly based (e.g. behavioral theory of the firm, strategic choice, dynamic capabilities), the multitude of the empirical investigations that contribute to the research stream that examines the factors that affect the magnitude of the impact of organizations on the evolution of their environment in general, and on the evolution of their technological environment in particular, leaves room for further research.

Moreover, within this study, we are entering a largely unexplored research area which concerns the determinants of a specific aspect of technological impact, namely the breadth of technological impact which refers to how broad is the organizations’ effect on technological evolution.

In the context of the contemporary rapid technological change, it is very important for inventors to seek and employ useful knowledge beyond their focal technological domains, in their effort to deal with technological problems of high complexity and to create breakthrough technological developments. Within this system of knowledge transfer, one can distinguish some organizations which systematically develop inventions that exert influence beyond their technological domains. Two typical examples of this kind of firm are the Xerox PARC and the Bell Laboratories, whose inventions are famous for their impact on technological developments in different technological fields (Rosenkopf and Nerkar 2001). From an evolutionary perspective, the technological inventions of those firms are very important since they serve as the connecting links across the different technological domains.

Therefore, the main objective of this study is to provide insight into the firm level factors that may affect a firm's ability to influence technological evolution and to create broadly useful technological inventions.

In addition, in the context of this thesis, we will examine if these two specific abilities can have an impact on the firm's financial performance. In particular, we will explore if the degree of a firm's technological influence is related to the degree of its financial success, adding to this specific but limited literature, and, moreover, we will analyze the relation between the breadth of impact and the financial performance, an issue on which the extant literature, to our knowledge, does not offer any empirical evidence.

Thus, the secondary objective of this study is to examine whether the firm's ability to develop influential inventions in general, or inventions of high breadth of impact in specific, can have a positive impact on its financial performance.

As far as the methodology is concerned, as we will show in more detail later, this study uses longitudinal (from 2003 to 2009), multi-industry (139 firms from biotechnology, pharmaceuticals, chemicals), multi-national (19 countries of origin) secondary data from two sources (EU Industrial R&D Investment Scoreboard and Derwent Innovation Index) and employs panel data econometric techniques to examine the hypotheses developed.

1.2 Contribution and Implications of the Study

Concerning the contribution of this thesis, we will first present the contribution which regards the examination of the factors that affect the breadth of impact and then we will proceed with the contribution which regards the magnitude of impact (i.e. the overall impact on technological evolution), since the breadth of impact is a rather understudied phenomenon and by far less examined that the magnitude of impact, and therefore, the contribution that pertains to the former is more important than the contribution that pertains to the latter.

In particular, and to the best of our knowledge, the extant literature provides us with only six studies that examine the determinants of the breadth of impact[1]. In comparison to those studies, we investigate the effect of four factors that have not been tested before, namely the extent to which a firm builds on scientific knowledge to develop new inventions, the degree of collaboration for the creation of new technological achievements, the degree to which a firm experiments with unexplored technological spaces and the size of a firm’s knowledge stock.

Moreover, we find three factors, which had been reported as significant predictors of breadth of impact in the previous studies, to be unrelated to it. More specifically, our results show that, contrary to the existing empirical findings, the degree to which a firm enters new-to-firm technological domains, the recency of the technologies that it incorporates into its inventions, and the degree to which it builds upon its prior research endeavors have no impact on the breadth of technological impact.

A final contribution of this thesis with respect to the determinants of the breadth of impact is that we not only include both of the dimensions and the measures of the

[1] These studies are presented analytically in subsection *2.4*.

breadth of impact that can be found in the literature but we also create a new one. In particular, the existing studies on this topic considered the breath of a firm's impact as either the extent of a firm's technological impact on industries others than its industry (*Beyond-Industry Impact*) or as the degree of concentration of the technological domains that the focal firm's inventions had an impact on (*Herfindahl Type Breadth*). Apart from those aspects of breadth of impact, we develop an additional one which considers the breadth of impact as the extent of a firm's technological impact on technological domains others than the ones in which it operates (*Beyond-Firm Impact*).

Concerning the contribution of this thesis with regard to the examination of the factors that affect the magnitude of impact, although there exists a non-negligible number of studies on this particular topic, this issue has been adequately examined[2].

A surprising result of this study, compared to the previous studies, is that it is not the broadness of the technological domains upon which the firm's inventions are based that positively affects the magnitude of impact but its opposite, that is, the narrowness of those technological domains, suggesting that it is the technological specialization that can lead to influential inventions and not the technological generalization. Another unexpected finding is that the degree to which a firm experiments with unexplored technological spaces affects negatively the magnitude of its technological impact, a finding which indicates that the disadvantages of the uncertainty that characterize the inventions that lack technological antecedents overcomes the advantages of their radicalness in the eyes of the inventors of the future technological inventions.

Moreover, in comparison to the extant literature, this study confirms that the recency of the technologies that are incorporated into a firm's inventions is a positive

[2] The main studies on this topic are presented analytically in subsection *2.4*.

predictor of the magnitude of impact. However and contrary to existing knowledge, this thesis does not confirm the positive effect of factors such as the size of a firm's knowledge stock, the degree of collaboration for the creation of new technological achievements, the degree to which a firm enters new-to-firm technological domains, the extent to which a firm builds on scientific knowledge to develop new inventions, and the degree to which it builds upon its prior research endeavors on the magnitude of impact.

The theoretical contribution of this study mainly lies on our effort to place the magnitude of technological impact into a general co-evolutionary framework, focusing on the organizations' effect on the evolution of their environment in general, and on the evolution of their technological environment in particular. The extant literature so far has used the magnitude of impact for the examination of various organizational phenomena (e.g. the nexus between breakthrough inventions and large established corporations, the linkage of basic science and technological impact, the structure of organizational knowledge bases, technological collaboration networks, R&D consortia and more) but not for the examination of the firm's impact on the evolution of its technological environment. The consideration of the magnitude of impact as a metric of the effect of each organization on the evolution of their technological environment and in combination with others studies that can quantify the effect of each organization on the evolution of its institutional environment (i.e. regulations, technological standards, cognition etc), can cause this study to lead to the emergence a new stream of research that will regard the quantification of the impact of each organization on the evolution of its environment (technological or institutional) and the factors that enhance or limit their ability to influence their environment.

Additionally, our study adds to the limited stream of research that considers the extent of a firm's technological influence not only a result of its technological competence and superiority but also a result of its social power. By the phrase “social power”, we mean the ability of a firm to influence various aspects of the institutional sphere such as regulations, legislations, cognitions, trends, and reputations.

Finally with regard to the factors that affect the magnitude and the breadth of technological impact, our study develops the more complete econometric model as compared to the econometrics models of the existing studies. In particular, we employ the biggest set of explanatory variables in our effort to explain better the determinants of the magnitude and the breadth of impact and to provide an econometric model with less biased estimators, based on the fact that the less the omitted variables that affect the response variable, the less biased the estimators.

The contribution of this study is not only limited to the determinants of the magnitude and the breadth of technological impact but also it adds to the stream of research that aims to assess the effect of magnitude of impact on financial performance and, more importantly, our study explores, for the first time to our knowledge, the effect of breadth of impact on financial performance.

As far as the implications of this study are concerned, we expect that the findings of this dissertation will be useful for various groups of people. Firstly, for those scholars who are generally interested either in the co-evolutionary relationship between organizations and their environment and especially in the organizations’ ability to affect their technological environment and to exert a broad influence on it or in the examination of the effect of the magnitude and the breadth of a firm’s impact on its financial performance.

Secondly, for the company managers who need to know the key characteristics that affect their firms’ ability to shape technological evolution and to exert a broad influence on technological change, since these abilities may prove to be very beneficial for them. In particular, the systematic adoption by other firms of a firm's technologies will probably increase the likelihood of firm's survival and growth and may result in better financial performance. In addition, managers with the knowledge of this study can focus on these firm-level competencies that increase the breadth of technological impact, in anticipation of the tangible benefits that may result from a potential innovative cooperation (e.g. joint venture) with firms from different technological domains, which, in turn, may lead to very successful innovative products.

Thirdly, for those investors who are looking for the optimum investments. Firms that have the ability to shape technological change may have higher survivability and a higher likelihood of achieving better financial performance in the long-run.

And finally, for those public policy makers who are interested in the formation of technologically powerful organizations with high ability to influence technological progress, for the benefit of their national innovation systems. If the public policy makers are cognizant of the effect on technological impact of specific factors on which they can exert an influence, they may attempt to increase or to decrease those factors aiming at the formation of a policy framework that favors the production of impactful technological achievements. For example, the policy makers could support and promote the firm – university cooperation or the R&D collaboration among firms or the R&D intensity by developing specific regulations (e.g. tax benefits) that favors this kind of corporate actions in anticipation of the emergence of technological powerful and influential firms at the global level.

The dissertation is organized into the following chapters:

- *Chapter 1*: Introduction

- *Chapter 2*: A review of the literature on the organizations' impact on their environment in general and on their technological environment in specific, and, in addition, a literature review on the breadth of technological impact followed by the conceptual framework of this study. Moreover, it includes the development of the hypotheses.
- *Chapter 3*: The detailed research methodology of the study. It includes the research setting, the sample, the measurement of the variables, the statistical models, and the descriptive statistics of the variables.
- *Chapter 4*: The results of our study. In addition, it includes an analysis about the effect of the firm's technological impact and breadth of impact on its financial performance.
- *Chapter 5*: The discussion of the research findings. The study concludes with its limitations, its managerial implications and the recommendations for further study.

CHAPTER 2

LITERATURE REVIEW & HYPOTHESES

2 LITERATURE REVIEW & HYPOTHESES

2.1 Co-evolution of Organizations and their Environment

Organizations shape and are shaped by their environment in a permanent, co-evolutionary manner. Generally, the co-evolution of a unit and its environment is a fundamental issue that concerns the entire spectrum of social sciences. The belief that co-evolutionary thinking approaches social phenomena in a realistic manner has attracted an increasing number of researchers, while co-evolution's complexity has caused the generation of different, scientifically interesting views of this concept.

Van Den Bergh and Stagl (2003) inform us that co-evolution was originally proposed in ecology to refer to the joint evolution of butterflies and flowering plants, denoting the fact that evolutionary changes in one species are a response to changes in other species with which it ecologically interacts.

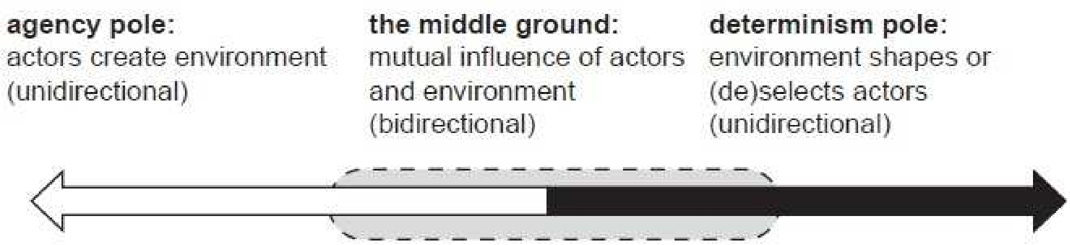
Management and organization studies are actively involved in the adoption and the development of co-evolutionary syllogism, a fact which is manifested through the application of this logic (explicitly or implicitly) in various co-evolving pairs either across different levels such as individual-firm, organization-meso/macro environment or group-organization or on the same level such as firm-firm, industry-industry and so forth (Baum and Singh 1994b, Rosenkopf and Tushman 1994, Lewin and Volberda 1999, Carney and

Gedajlovic, 2002, Flier et al. 2003, Rodrigues and Child 2003, Volberda and Lewin 2003, Dieleman and Sachs 2008, Cantwell et al. 2009, Murman 2013).

Within the realms of management and organization theories, Lewin and Volberda (1999) defined co-evolution as the joint outcome of managerial intentionality, environment, and institutional effects. According to them, "change can be recursive and need not be an outcome of either managerial adaptation or environmental selection but rather the joint outcome of managerial intentionality and environmental effects" (Ibid., p. 526).

Actually, co-evolutionary reasoning occupies the middle point on the imaginary line between the deterministic and the volitional pole (Wohlgezogen and Hirsch 2009). Two indicative examples of research streams on the basis of this polarization are the adaptation-selection debate grounded in the organization theories (Greenwood and Hinings 1996, Lewin and Volberda 1999, Volberda and Lewin 1999) and the agency-structure debate grounded in the institutional theories of organization (Hodgson 2004, Heugens and Lander 2009, Bailey and Barley 2011). On the one hand, the adaptation and the agency sides both emphasize organization's power, influential activity and idiosyncrasy, whereas, on the other hand, the selection and the structure sides give extra weight to organization's inertia, passivity and isomorphism.

Figure 2.1 The Agency–Determinism Continuum



Source: Wohlgezogen and Hirsch (2009)

The preponderance of co-evolution's logic over the poles' logic has been underlined by many authors. Pacheco et al. (2010) claimed that institutions may constrain actors' actions but they do not determine them. Moreover, Barley and Tolbert (1997) argued that although institutions set bounds on rationality by restricting the alternatives we perceive, individuals and organizations, through choice and action, can deliberately modify, create and even eliminate them. In the same vein, Carney and Gedajlovic (2002) highlighted that actors are often capable of influencing institutional arrangements, but are nevertheless subject to constraints that limit their range of feasible and conceivable action. In addition, Murmann (2013, p. 61) stressed that "what sets co-evolutionary theory apart from standard evolutionary explanations is that causality does not only run from the environment to the evolving entity but it also runs from the entity to the environment" and he highlighted the case of the Internet where firms are not only dramatically affected by Internet technologies, but also certain firms like Microsoft, Apple, or Google have a huge impact on how Internet technologies evolve. Furthermore, Hodgson (2013) prompted scholars for some reconciliation between the organizational ecology approach (Hannan and Freeman 1993), with its stress on selection and highly limited adaptation, and the business strategy literature, with its stress on the need for adaptability. Finally, apart from its theoretical importance, the co-evolutionary view, according to Baum and Singh (1994a), can potentially add value to more practical domains such as public policy or regulation.

A constant complaint of the co-evolutionary proponents is that studies of how organizations influence their environment are rarer than studies of how they adapt to it

(Stern and Barley 1996). Hitherto, the extant literature primarily has focused on how and why organizations tend to become isomorphic with their environments, whereas questions of how organizations systematically influence their environments are studied less (Baum and Singh 1994b, Lewin and Volberda 1999). Attempting to explain this gap, Dieleman and Sachs (2008) stressed that scholars assume often that corporations are too insignificant to have an impact on their environment, while Barley and Tolbert (1997) underscored the difficulties on detecting and collecting data for documenting the effect of an organization's activities' on its context.

Despite the lack of empirical investigations on the influence of actors on their environment, there exist some theories that acknowledge and build on the shaping aspect of organizations. The extent to which each theory focuses on the organizations' capacity to mold their social landscape varies substantially. For example, the behavioral theory of the firm, although primarily views organizations as capable of changing their goals, focus of attention, and search procedures, secondarily suggests that firms can affect the external environment in which they operate (Lewin et al. 2004). Furthermore, both evolutionary economics and dynamic capabilities approaches accept the organizations' power to influence their environment. Evolutionary economics focuses on firms as vehicles of innovation and drivers of change at the industry level (Lewin et al. 2004) while the firms' capability of manipulating their external context is a constituent part of dynamic capabilities (Teece 2009). In addition, the strategic choice approach contends explicitly that firms, on the basis of managerial intentionality, have the ability to reshape their environment rather than simply being passive recipients of environmental forces (Flier et al. 2003, Lewin et al. 2004).

The viewpoint that organizations constantly shape their environment gains the significance it deserves within the co-evolutionary framework provided by Lewin and

Volberda (Lewin and Volberda 1999, Volberda and Lewin 2003). According to them, co-evolution is the "joint outcome of managerial intentionality, environment, and institutional effects" (Lewin and Volberda 1999, p. 526) and any empirical co-evolutionary research must satisfy certain requirements such as longitudinal time series of microstate adaptation events, modeling multidirectional causalities between micro and macro co-evolution, and analyzing for mutual, simultaneous, lagged and nested effects.

A representative approach on the shaping role of organizations is the concept of institutional entrepreneurship. DiMaggio (1988) introduced the notion of institutional entrepreneurship trying to explain 'institutionalization' as a continuous process (Garud et al. 2002) and to analyze how actors can contribute to changing institutions, despite pressures towards isomorphism (Battilana et al. 2009). Institutional entrepreneurs are organizations or individuals who leverage resources to create new or transform existing institutions (Battilana et al. 2009), driven by a wide range of motivations and operating in a variety of contexts (Pacheco et al. 2010). Institutional entrepreneurship, as embedded within the structure–agency debate, is a typical co-evolutionary concept where institutions constrain and enable, but do not determine, the choices of actors (Battilana et al. 2009), who are, in turn, the agents of institutional change (Pacheco et al. 2010).

The existing literature provides us with several studies that demonstrate the influence of an actor on all three aspects of institutions, namely, the regulative, the normative, and the cognitive (Hoffman 1999). Regarding the most commonly studied aspect, the regulative, Garud et al. (2002) showed how Sun Microsystems' shaped a common technological standard by sponsoring its Java technology, and Bekkers et al. (2002) investigated the role of intellectual property rights' (IPRs) ownership in shaping the GSM (global system for mobile communications) standard. Additionally, Schuler et al. (2002) emphasized the importance for firms to obtain and to maintain access to those who formulate public policy in an endeavor to gain influence over the legislative processes. Moreover, Dieleman and Sachs (2008) examined whether large family groups in emerging economies can shape institutions to their advantage by influencing politicians, while Rodrigues and Child (2003) studied the case of a major Brazilian telecommunications company that affected the changes in the rules of market competition. Cantwell et al. (2009) and Child and Tsai (2005) both stressed the political activities, such as lobbying and public relations, that multinational companies take in pursuit of shaping favorable regulations. Regarding the normative aspect, Huygens et al. (2001) demonstrated how firms in music industry developed a new competitive regime by introducing new practices that replaced the existing business models and Cantwell et al. (2009) researched how multinational enterprises influence local organizational routines by transferring their best practices across countries. Regarding the cognitive aspect, Munir and Phillips (2005) illustrated the strategies that Kodak followed for the transformation of photography from a highly specialized activity to one that became an integral part of everyday life, and Rindova and Petkova (2007) described firms' impact on customers' cognitive and emotional state through the specific functional, symbolic, and aesthetic properties of firms' technological innovations. Finally, Kaplan and Tripsas (2008) underlined how producers of new technologies can shape actors' cognition through media or though the interactive process of joint learning and experimentation.

2.2 Organizations' Impact on Technological Evolution

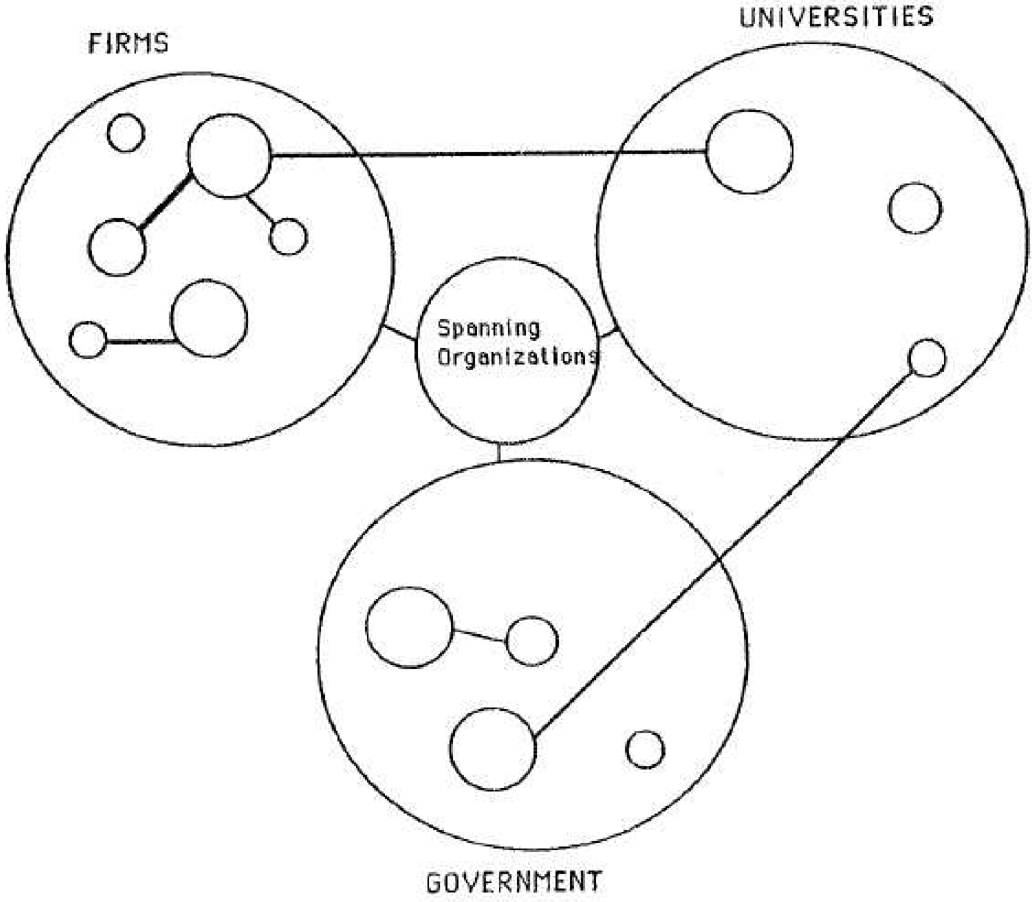
The co-evolutionary theorizing can be preserved also when we consider a certain aspect of the general environment, the technological environment. Within the co-evoluton of organizations and technology, the causality does not only run from the technological environment to the evolving organizations (i.e. all the existing knowledge upon which an organization is based to develop new technological inventions) but it also runs from the organization to the technological environment (i.e. the new technological inventions developed by the organization) (Murmann 2013).

Organizations and technology are two co-evolutionary partners that interacting from different levels, that is to say, at the micro level (organization) and at the macro level of analysis (technology). Each inventing organization as being nested with a certain technological environment, "receives" the technical knowledge from that environment, and through its R&D operations, achieves to develop a new technological invention. This technological invention is not only a step in the organization's evolutionary path, but at the same time, it is a step in the evolutionary path of the technological environment.

In the extant literature, the most influential model of the co-evolution of organizations and technology is based on the interaction between technological communities and technological progress in the context of technology cycles (Rosenkopf and Tushman 1994). In particular, this model analyzes how structural changes in technological communities drive and are driven by technological evolution, during the four components of technology cycle (i.e. technological discontinuity, era of ferment, dominant design, and era of incremental change).

Rosenkopf and Tushman (1994) employed the technological community (*Figure 2.2*) as the unit of analysis for the study of technological change and they defined it as the set of organizations that are stakeholders for a particular technology (this set of organizations can include suppliers, manufacturers, user groups, governmental agencies, standards bodies, and professional associations).

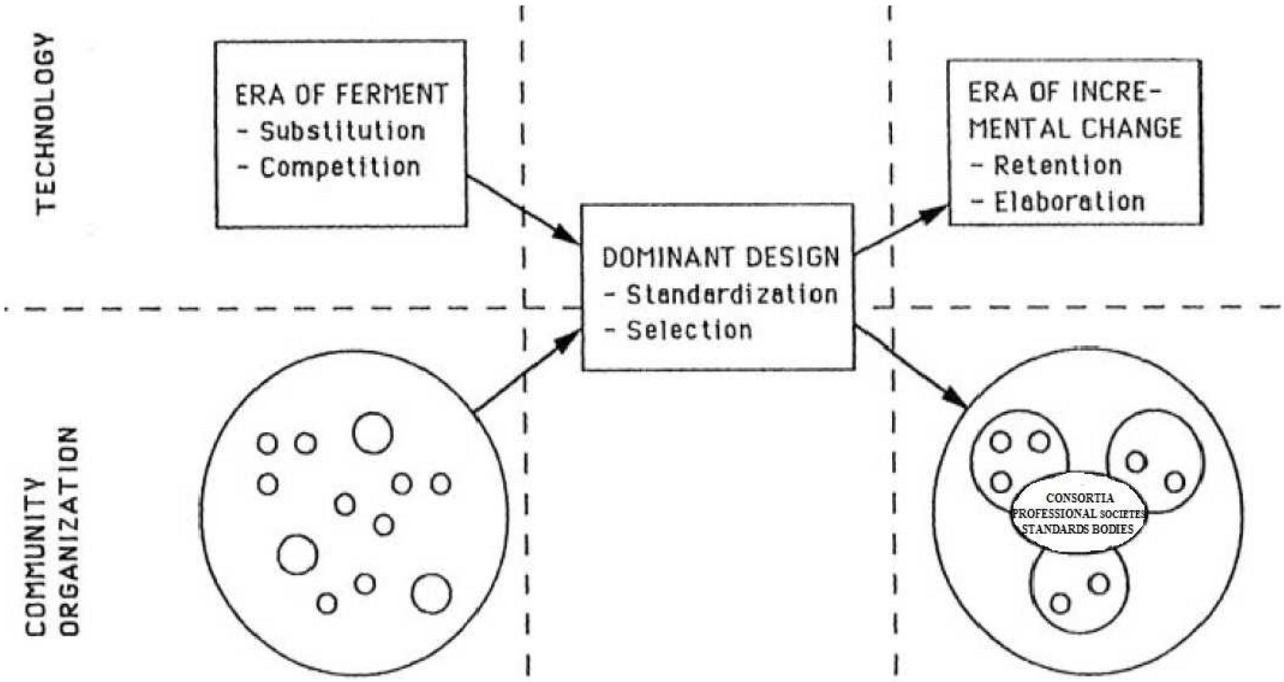
Figure 2.2 Technological Community



Source: Rosenkopf and Tushman (1994)

The main thesis of their study is that when a new technological discontinuity is introduced to the technological community, new actors and new interorganizational linkages emerge within the community. Then, the era of ferment follows, during which, interorganizational activity focuses on developing markets and matching technology to these markets, as technology stakeholders attempt to influence the outcome of technological competition in their favor. It is important to note that according to the authors “organizations must develop not only technical competence, but also interorganizational network skills to forge alliances in order to shape critical dimensions of merit and critical industry problems” (p. 415). The selection of a dominant design settles technological competition and brings the era of ferment to an end. The selection of dominant design occurs through the determination of appropriate dimensions of merit for the new technology and emerges not only from technical logic, but also from “a negotiated logic enlivened by actors with interests in competing technical regimes” (p. 419). Finally, during the era of incremental change, the primary function of the community is the elaboration of the standard. Technical uncertainty decreases and the nature of technical change shifts from variation to incremental change. Technical clarity and convergence on a set of technical parameters permit firms to design standardized and interchangeable parts and to optimize organizational processes for volume and efficiency and focuses on process-based improvements that reduce cost and uncertainty community (see *Figure 2.3* for a schematic representation of the co-evolution of technology and organizations).

Figure 2.3 Co-evolution of Technology and Organization



Source: Rosenkopf and Tushman (1994)

It is important to underline the emphasis that the authors placed on the non-technological capabilities (i.e. inter-organizational network skills) that organizations must develop in order their technological achievements to be accepted by the technological community. These capabilities are of great importance especially during selection of dominant design where “social logic drives technical progress as suppliers, customers, or governments react to the uncertainty and inefficiencies associated with eras of ferment” (Ibid., p. 419). The same attention on the social dynamics underpinning dominance has been paid by Tushman and Murmann (1998) and by Munir and Jones (2004), who also emphasized the various non-technological, sociopolitical factors that are understood to play a critical role in the success or failure of designs.

Another significant point in this study is the proposed construct of power. Power refers to the differences in the actors’ (i.e. organizations’) abilities to shape and influence the paths of technological change. “Firms with technology strength or market strength control resources that other organizations need, and power accrues to these firms as others need access to new technology or to the market” ”(Ibid., p. 412). An important feature of this power is the organization’s social or political status within the community, such as regulatory bodies, standards bodies, or other associations, which can strongly influence technological progress. The power of any actor to shape technological change socially can grow or diminish over time and evolves in parallel with the technology cycle. In particular, during the era of ferment the emergence of new actors and new linkages leads to a redistribution of power among these players, while during the era of incremental change power accrues to the dominant coalition of actors.

Although this study acknowledged power as the organizations' ability to shape technological evolution, it did not investigate specifically the factors that can reinforce it or weaken it. The main interest of Rosenkopf and Tushman (1994) with regard to power was to investigate the concentration of power during the era of ferment and during the era of incremental change. In particular, their major proposition regarding power was that "The distribution of power becomes less concentrated during the era of ferment and more concentrated during the era of incremental change" (ibid, p.422). Therefore, they examined power at the aggregate level of analysis (i.e. at the technological community level of analysis) and not at the firm level of analysis.

2.3 Breadth of Technological Impact

One of the main characteristics of modern technology is the high degree of interconnectivity among the different technological domains. The high levels of technological complexity and the demand of the intense competition for quick and breakthrough technological solutions has led inventors to seek useful knowledge beyond their focal technological domains. Miller et al. (2007) trying to emphasize the importance of the interaction among technologies mentioned that inventors may not be able to design a simple appliance or even a toy that will have the features that customers want, if they are not familiar with the latest discoveries in microelectronics.

Within this system of knowledge transfer, there exist firms which mainly play the role of "technological seeders", that is to say, they systematically develop inventions that exert influence beyond their focal technological domain. This particular dimension of technological impact, the breadth of impact, is a very important issue which deserves to be examined separately, since, as technology evolves, the interconnectivity and the interactions among different technological fields are strongly based on these inventions whose influence goes beyond the boundaries of their technological domain.

For example, Rosenkopf and Nerkar (2001) stressed that Xerox PARC and Bell Laboratories are famous for the development of technologies with implications far beyond their traditional markets. In addition, Steve Kerr, the vice president of Corporate Leadership Development and chief learning officer of General Electric Corporation in 1997, emphasized the importance of moving ideas around diverse businesses that don't have a lot in common. In particular, he stated that "*A breakthrough in GE's Medical Systems business, with relatively little modification, led to a method by which an aircraft engine can transmit continuous information about blade speed, engine heat and other relevant data about its in-flight performance well in advance of any possible safety situation. This innovation, in turn, catalyzed an important new development with respect to a self-monitoring system for use with heart pacemakers.*" (Miller et al. 2007, p. 308). Moreover, Banerjee and Cole (2010) stressed that some technologies such as James Watt's steam engine moved successfully across broad technological fields as widely divergent as water pumping in the coal mining industry to propelling locomotives in the railroad industry, while Hall et al. (2001) showed that, at the industry level of analysis, information and communication technologies (ICT) can be characterized as general-purpose technologies because of their widespread technological impact.

The breadth of technological impact is conceptually closely related to the tension between exploitation and exploration (or between local and distant knowledge), which is a major issue regarding technological innovations. As Miller et al. (2007) put it, technological development requires the synthesis of two divergent tasks. On the one hand, a firm must focus on a bounded set of technological domains, in order to be able to produce technological inventions by recombining knowledge elements from a familiar space or refining an existing knowledge combination. This local search (i.e. exploitation) will help the firm to create valuable and commercially viable products. On the other hand, a firm must continually search for diverse and distant knowledge elements that may lead the firm to produce path-breaking innovations. This distant search (i.e. exploration), although can be costly and may never lead to viable or valuable innovations, is fundamental to the development of new technological capabilities and for enabling "the transition from an entrenched set of techniques and designs to a new technological paradigm" that can be truly innovative (Miller et al. 2007, p. 309).

Therefore, these explorative firms, driven by their motivation to create groundbreaking technological solutions based on the combination of the old-to-firm knowledge components with the new-to-firm knowledge components, seek for inventions from different technological fields that encompass pieces of technical knowledge which are characterized not only by their high value in solving technological problems but also by their generality in utilization from technologies from different technological fields. So, knowledge components that provide efficient solutions to certain technological problems and, at the same time, they can be used in different technological domains are the top choices for the explorative firms, in their quest for radical solutions to their technological challenges.

Valentini and Di Guardo (2012) suggested that the ability of firms to create technologies with powerful breadth of impact depends on their capability to master the logics of induction and deduction, to come to general conclusions from particular instances, to establish general law from the observation of particular cases and, consequently, to generate general solutions to technical problems. Technologies that encompass knowledge components that provide general solutions can be utilized and be useful in different technological fields.

2.4 Factors Affecting the Magnitude and the Breadth of Impact

In an increasingly knowledge-based and competitive economy, the survival – let alone the success – of a firm is strongly dependent on its ability to be “actively involved” in the shaping of technological change. By the phrase “actively involved”, we mean not only to be able to detect, to understand and to assimilate the new technological achievements but also to be able to create new and useful technological knowledge. The absence of the latter ability will jeopardize the firm’s own survival because this firm will be always one step backwards in the “technological race” from its competitors who have the ability to develop new technical knowledge and to determine the technological route.

So, we can argue that the success of a firm is strictly based on the success of their technological inventions, that is to say, on the extent to which their technological developments have been prevailed over other rival technological developments or on the extent to which their technological inventions are perceived as worthwhile by the organizations that directly determine whether it is used or not. Thus, although firms generally are driven by “profit-seeking” or “growth-seeking” rather than ‘‘impact seeking’’ concerns, it is not unreasonable to claim that the developing of impactful technologies is a crucial element in the firms’ struggle for survival and success (Banerjee and Cole 2010). Therefore, the understanding the factors that shape organizations’ ability to produce influential technological knowledge is a very important issue in organizational studies (Valentini and Di Guardo 2012).

Turning to our study, we remind readers that the main focus of this thesis is to examine the **firm level factors** that can limit or enhance the power of an organization to shape technological evolution in general and to exert a broad influence on technological change in specific. Conceptualizing this relation, we categorize firm characteristics into two groups according to whether a firm characteristic has an effect on the technological or on the non-technological capabilities of the firm. In other words, according to whether a firm level factor is influencing the firm's ability to shape technological change via its technological superiority or via its social power and influence.

To this end, on the one hand, we examine firm characteristics such as R&D investment, technological collaboration, technological diversity, R&D scientific basis, and others, that are considered to have an impact only on the firm's technological performance and, via that, on the establishment of firm's technological dominance. On the other hand, the firm size is a firm characteristic that apart from its effect on technological performance[3], it can have an impact on the magnitude and the breadth of impact by influencing the firm's "social power". In this context, by the phrase "social power", we mean the firms' ability to exert strong influence on their technological community to "accept" their technological developments, either by using their prestige and their reputation or by financing industry associations and lobbyism, by promoting technological standards, and by putting pressure on governments to regulate to their benefit. Moreover, the recent technological activity of a firm (i.e. the knowledge stock upon which a firm builds its new technological inventions) can enhance or limit the firm's technological reputation and its attractiveness to other firms apart from its effect on the firm's technological competence.

At this point we must make a brief reference to how we measure the firm's ability to shape technological change in terms of magnitude and in terms of breadth. A firm is considered to affect technological progress when its technological achievements are

[3] Variations in firm size can affect the extent of bureaucracy and inertia within an organization as well as the economies of scale and scope. All of these factors can impact the organization's capability to develop numerous and influential technological achievements, See the developments of the Hypothesis 9 is subsection 2.5.10 for a more analytical examination.

affecting future technological achievements. Technological achievements can be captured by measuring the firm's patented inventions, while the ***magnitude of impact*** on subsequent technological achievement can be measured by the number of citations that these patents receive by subsequent patented inventions. For the ***breadth of impact***, we use three different measures based on the patent citations received. Briefly at this point, we mention that the first measure focuses on the firm's impact on technological areas that belong to industries different from the firm's industry (we call this type of measure *Beyond-Industry Impact*); the second focuses on the impact on technological areas different from the areas that the focal firm operates in, regardless of the industry to which these technological areas belong (we call this type of measure *Beyond-Firm Impact*); and the third measure – which we call it *Herfindahl Type Breadth* – is build on the basis of the Herfindahl Index and it focuses on how concentrated or how dispersed are the technological fields of the affected patented inventions (see subsection *3.3.1.1* and *3.3.1.2* for the measurement of the magnitude and the breadth of technological impact, respectively).

Therefore, in the conceptualization of the relation between the firm level factors and the firm's ability to shape technological change, we have to include also a number of patent level characteristics that may affect our dependent variable of interest.

It is important to emphasize that firms are aiming and endeavoring to develop successful technological developments that will prevail over rival technologies ***and not*** to develop patented inventions that will receive a substantial high number of patent citations. According to this logic, we actually provide an ***ex-post*** explanation for why some firms affects technological evolution, meaning by the word ***ex-post***, that the magnitude and the breadth of technological impact, as measured by patent citations, cannot be considered as a firm's strategic choice but only as the outcome of their effort to prevail technologically and, so, the R&D managers and the inventors are probably not aware of the causal relationship between the examined factors and the number of patent citations received[4]. Studies such as this one actually attempt to explain, in retrospect, which were the main determinants of the magnitude and of the breadth of technological impact. For example, when a firm decides to invest more in R&D or to collaborate with other firms to develop new inventions, its goal is to produce technological superior inventions that can solve important technological problems and not to develop patented inventions that receive numerous citations. Nevertheless, the number of citations that the patents of a firm receive can indeed be used as a proxy of the firm's ability to superiority of the technological inventions.

Figure 2.4 illustrates our conceptual framework regarding the factors that affect firm's ability to develop inventions of high magnitude and of high breadth of impact. Except for the firm level factors, our conceptual framework is composed of three more types of factors, namely, the patent level, the industry level, and the country level factors, all of which can have a significant impact on the number of patent citations, and so, we have to account for them.

More specifically, with regard to the firm level factors examined in this study, we placed them into two categories: factors that affect magnitude and breadth of impact by contributing to the technological superiority of a firm and factors that affect magnitude and breadth of impact by contributing to the social power of a firm[5]. Except for the factors of these two categories, we control also for profitability (category *Other Firm*

---

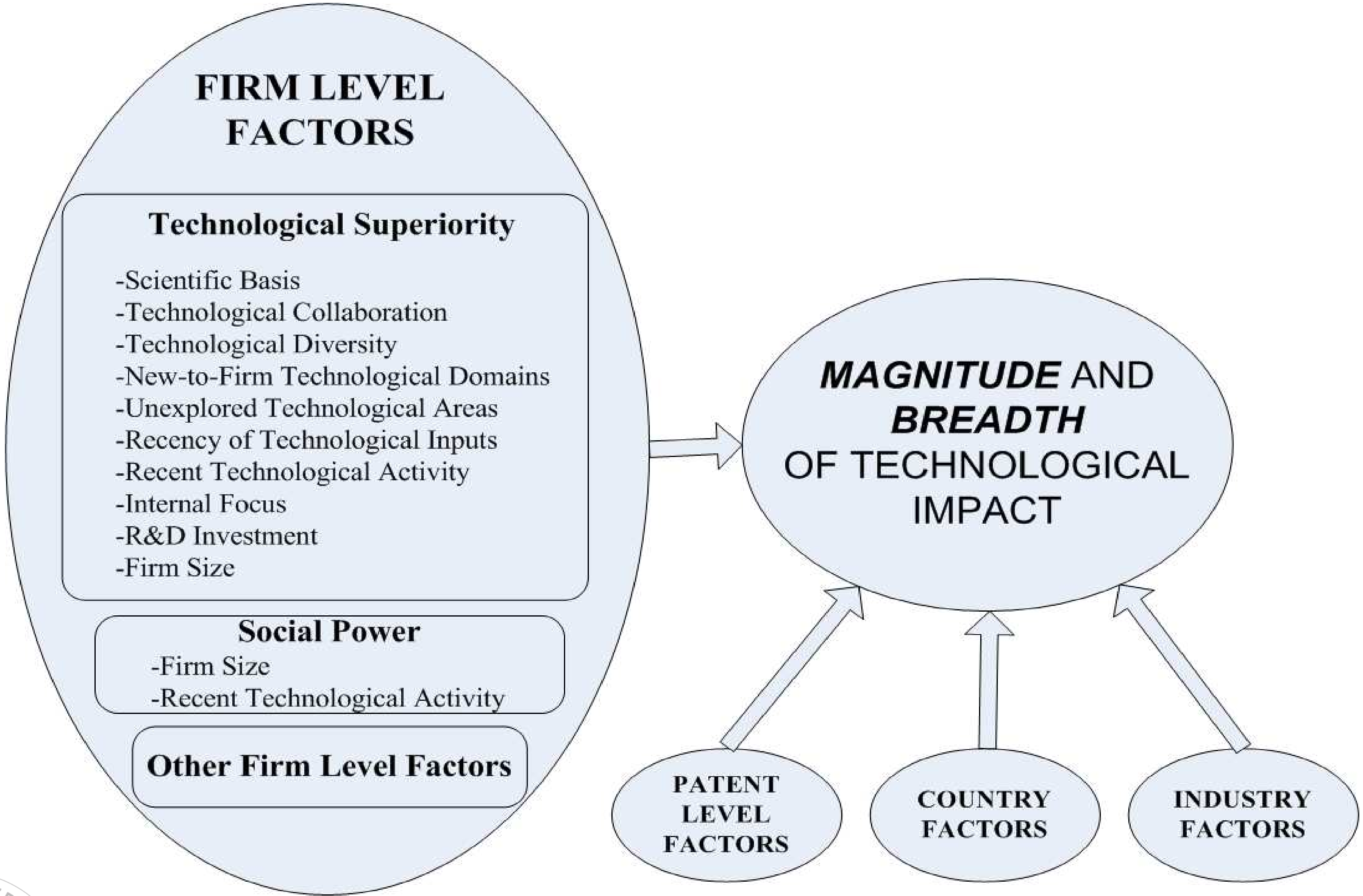
[4] This is especially true for the breadth of impact. The extent of the breadth of impact cannot be easily considered as choice of the management team but it can be easily considered as an unplanned result.

[5] *Firm Size* and *Recent Technological Activity* can influence magnitude and breadth of impact in both ways. So, in the schematic representation of the conceptual framework in *Figure 2.4,* they appear in both categories.

Level Factors), which is a firm level factor that can influence the number of patent citations.

Although our conceptual framework attempts to incorporate the full set of the factors that can affect the number of patent citations and even if we include the explanatory variable *Recent Technological Activity*, which has been used by many studies as a control for the unobserved heterogeneity among the sample firms (Dushnitsky and Lenox 2005, Srivastava and Gnyawali 2011), there may still exist factors that could influence our dependent variable and are not completely controlled in our model. For example, mergers and acquisitions (M&A) have been shown to affect the number of patent citations received (Valentini 2012, Valentini and Di Guardo 2012). Although we couldn’t collect data on M&A within the context of this study, we believe that a large part of this effect will be controlled by the variable *Firm* Size, based on the assumption that each merger or acquisition is accompanied by an analogous increase in the size of the firm.

Figure 2.4 Conceptual Framework: Factors affecting Firm's Magnitude and Breadth of Technological Impact



With regard to the extant literature on the factors that affect the magnitude of technological impact, there exists a non-negligible number of studies on this particular topic, but no one can argue that this issue has been adequately examined. In *Table 2.1*, we present the main studies that examine the determinants of magnitude of technological impact, accompanied by their main area of interest, by the main determinants examined, and by number of independent variables (control variables included) that are contained in their econometric model.

These studies use patent citations as an indicator of the magnitude of technological impact in an effort to assess firm technological performance either in terms of developing usefulness and valuable inventions (e.g. Fleming and Sorenson 2001, George et al. 2008, Yayavaram and Ahuja 2008, Joshi and Nerkar 2011) or in terms of developing radical and breakthrough inventions (Ahuja and Lampert 2001, Schoenmakers and Duysters 2010, Srivastava and Gnyawali 2011)[6].

There is great variety in the theoretical approaches employed by the authors of these studies, as clearly shown in the column *Main Area of Interest* of *Table 2.1*. Magnitude of impact has been used in the examination of various organizational phenomena such as the nexus between breakthrough inventions and large established corporations (Ahuja and Lampert 2001), the uncertainty in technological search (Fleming 2001), the temporal knowledge exploration (Nerkar 2003), the linkage of basic science and technological impact (Sorenson and Fleming 2004), the structure of organizational

[6] It is important to note that because the authors of these studies use patent citations as a proxy for radical or breakthrough inventions, whenever they refer to radical inventions they usually refer to the patents that receive numerous patent citations. However, in our study the term radical or breakthrough inventions refers to those inventions whose novelty of the combinations of their knowledge components is so fundamentally different from previous inventions that they can serve as the basis of new technological trajectories and paradigms. So, in our study radicality characterize the novelty of the technological knowledge and not the amount of citations received. Nevertheless, as will be shown later in the development of the hypotheses, we use the logical assumption that when a patented invention is radical in terms of technological knowledge is likely to be cited more heavily as compared to an invention that is characterized as incremental.

knowledge bases (Yayavaram and Ahuja 2008), technological collaboration networks (Operti and Carnabuci 2014), R&D consortia (Joshi and Nerkar 2011) etc..

However, none of these studies placed the magnitude of technological impact into a general co-evolutionary framework, focusing on the organizations' effect on the evolution of their environment in general, and on the evolution of their technological environment in particular. Considering the magnitude of impact as a metric of the effect of each organization on the evolution of their technological environment and in combination with other studies that quantify the effect of each organization on the evolution of its institutional environment (i.e. regulations, technological standards, cognition etc), this study may lead to the emergence a new stream of research that will regard the quantification of the impact of each organization on the evolution of its environment (technological or institutional) and the factors that enhance or limit their ability to influence their environment.

Moreover, as will be shown later in *Chapter 3*, the number of the independent variables (including the control variables) of our econometric models with regard to magnitude of impact is 21[7] which is the highest number among the studies of *Table 2.1* as depicted in column *Number of Independent Variables*[8]. Consequently, we can argue that we not only examine the largest set of explanatory variables but we probably provide the less biased estimators, since the effect of the error of the unobserved omitted variables is reduced (the less the omitted variables that affect the response variable, the less biased the estimators).

Table 2.1 Empirical Studies that Use Magnitude of Impact as Response Variable

[7] Including the country and industry dummies variables.
[8] Average number: 10.82, max: 18 (Miller et al. 2007, Operti and Carnabuci 2014), min: 5 (Schoenmakers and Duysters 2010).

	Study	Main Area of Interest	Determinants of Magnitude of Impact	Number of Independent Variables (Controls Included)
1.	Ahuja and Lampert (2001)	The Nexus Between Breakthrough Inventions and Large Established Corporations	1. Novel Technologies (Technologies in which the Firm Lacks Prior Experience) 2. Emerging Technologies that are Recent or Newly Developed in the Industry 3. The Extent to which an Invention is Based on Previous Inventions	9
2.	Fleming (2001)	Uncertainty in Technological Search and Technological Impact	1. Recombination of Familiar Knowledge Components 2. Refinement of Familiar Combinations 3. Cumulative Use of a Combination (Technological Exhaustion)	10
3.	Fleming and Sorenson (2001)	Complex Adaptive Systems Theory and Technological Impact	1. Number of Knowledge Components 2. Degree of Interdependence Between Knowledge Components (the Functional Sensitivity of an Invention to Changes in these Constituent Components)	9
4.	Gittelman and Kogut (2003)	The Conflicting Logics between Science and Innovation	1. Science Intensity of Firm's Inventions 2. Firm Average Citations to Publications 3. Firm Publications 4. Firm Co-publications	14
5.	Nerkar (2003)	Temporal Exploration and Technological Impact	1. Recency of Technological Inputs 2. Time spread of Technological Inputs	12
6.	Sorenson and Fleming (2004)	The Linkage of Basic Science and Technological Impact	1. Science Intensity of Firm's Inventions	10

	Study	Main Area of Interest	Determinants of Magnitude of Impact	Number of Independent Variables (Controls Included)
7.	Sapsalis et al. (2006)	Academic versus Industry Patenting	1. Science Intensity of Firm's Inventions 2. Technological Collaboration 3. The Extent to which an Invention is Based on Previous Inventions	12
8.	Hoetker and Agarwal (2007)	The Diffusion of Knowledge After a Firm's Exit from an Industry	1. Number of Inventors 2. Range of Technologies Combined 3. Firm Exit 4. Firm Age 5. Internal focus	10
9.	Miller et al. (2007)	Knowledge Exploration across Divisions in a Diversified Firm and Technological Impact	1. Intradivisional Knowledge 2. Extraorganizational Knowledge 3. Interdivisional Knowledge	18
10.	Nagaoka (2007)	R&D Management and Technological Impact	1. Science Intensity of Firm's Patents 2. Recency of Technological Inputs	9
11.	George et al. (2008)	Reliance on Prior Innovations and Technological Impact	1. The Degree to which Prior Innovations within the Same Technical Domain Help Generate New Innovations 2. The Depth of Technical Expertise in a Given Technical Area 3. Technological Diversity	9

	Study	Main Area of Interest	Determinants of Magnitude of Impact	Number of Independent Variables (Controls Included)
12.	Yayavaram and Ahuja (2008)	Structure of Organizational Knowledge Bases and Technological Impact	1. Variations in Coupling Patterns Between Knowledge Elements (Fully Decomposable - Knowledge Elements with no Significant Ties - Discernable but Connected - Couplings are Pervasively Distributed)	9
13.	Schoenmakers and Duysters (2010)	The Specific Nature of Radical Inventions	1. Technological Diversity 2. Use of Mature vs Emerging Technologies 3. The Extent to which an Invention is Based on Previous Inventions	5
14.	Joshi and Nerkar (2011)	R&D Consortia and Technological Impact	1. Participation in Patents Developed by R&D Consortia	8
15.	Srivastava and Gnyawali (2011)	Quality and Diversity of Portfolio Technological Resources and Technological Impact	1. Quality and Diversity of the Portfolio of Technological Relational Resources 2. Technological Diversity 3. Technological Strength	15
16.	Nemet and Johnson (2012)	Transfer of New Knowledge from one Technological Domain to another and Technological Impact	1.External Knowledge (Citations to Different Technological Domains)	7

	Study	Main Area of Interest	Determinants of Magnitude of Impact	Number of Independent Variables (Controls Included)
17.	Operti and Carnabuci (2014)	Technological Collaboration Networks and Technological Impact	1. Spillover Network Munificence (How Munificent is a Firm's Spillover Network) 2. Structural Holes (the Degree of Connectivity –or the Lack of it– Between a Firm's "source" Firms)	18

With regard to the extant literature on the factors that affect the breadth of technological impact, it is not difficult to figure out that this research area is seriously understudied. To our knowledge, there exist only six studies that examine the determinants of the breadth of impact. *Table 2.2* displays these six studies together with their main area of interest, the type of measure the authors used for the breadth of impact, the main determinants examined, and the number of independent variables (control variables included) that are contained in their econometric model.

The column *Measurement of Breadth of Impact* in *Table 2.2* shows that four studies employed the *Herfindahl Type Breadth* as a measure of the breadth of impact while two studies employed the *Beyond-Industry Impact*. However, none of these studies used the *Beyond-Firm Impact*, a fact which makes our thesis to be the first study that it does.

Moreover, as was the case with the magnitude of impact concerning the set of explanatory variables of our econometric models compared with those of the extant literature, the number of our independent variables (a total of 21 including the control variable) is the highest among the studies of *Table 2.2* as depicted in column *Number of*

Independent Variable[9]. Therefore, not only, our model is the most complete econometric model but also it can produce the less biased estimators, since it suffering the least from the error of the unobserved omitted variables.

Table 2.2 Empirical Studies that Use Breadth of Impact as Response Variable

	Study	**Main Area of Interest**	**Type of Measure of Breadth of Impact**	**Determinants of Breadth of Impact**	**Number of Independent Variables (Controls Included)**
1.	Rosenkopf and Nerkar (2001)	Organizational and Technological Boundary-Spanning Exploration	Beyond-Industry Impact	1. Exploration within Organizational Boundaries 2. Exploration within Technological Boundaries	11
2.	Argyres and Silverman (2004)	The Firm's Choice to Operate a Centralized or Decentralized R&D Structure and the Type of Innovation it Produces	Herfindahl Type Breadth	Centralized / Decentralized R&D Activity	13
3.	Miller et al. (2007)	Knowledge Exploration across Divisions in a Diversified Firm and Technological Impact	Beyond-Industry Impact	1. Intradivisional Knowledge 2. Extraorganizational Knowledge 3. Interdivisional Knowledge	18

9 Average number: 11.66, max: 18 (Miller et al. 2007, Operti and Carnabuci 2014), min: 9 (Banerjee and Cole 2010, Valentini and Di Guardo 2012).

	Study	Main Area of Interest	Type of Measure of Breadth of Impact	Determinants of Breadth of Impact	Number of Independent Variables (Controls Included)
4.	Banerjee and Cole (2010)	Knowledge Exploration Choices and Breadth of Impact	Herfindahl Type Breadth	1. Entering new Application Domains 2.Sourcing Diverse Areas of Knowledge	9
5.	Valentini (2012)	The Effect of M&A on the Patenting Quantity and Quality	Herfindahl Type Breadth	Mergers and Acquisitions (M&A)	10
6.	Valentini and Di Guardo (2012)	Mergers and Acquisitions (M&A) and the Profile of Merging Firms' Inventive Activity	Herfindahl Type Breadth	1. Diversity of the Merging Firms' Knowledge Bases 2. Market Relatedness of an Acquirer and a Target	9

2.5 Hypotheses

2.5.1 Hypothesis 1 - Scientific Basis

Magnitude of Impact: Scientific basis refers to the extent to which firms build on scientific knowledge while trying to resolve technological difficulties and to produce new inventions (Operti and Carnabuci 2014). Although scientific knowledge is created for scientific purposes and not for commercial ones, firms systematically search within scientific literature and endeavor to apply that knowledge to the development of new technologies (Gittelman and Kogut 2003).

The use of scientific knowledge within firms' inventions implies the transfer of pieces of knowledge from a theoretical level to a practical level. This transfer of scientific knowledge in the sphere of technological inventions may cause the development of a radical and pioneer technology. And although, radical inventions are not identical to impactful inventions, they are likely to be, in general, more influential in shaping technological change than the incremental inventions (and, thus, radical patented inventions are likely to receive more citations from future patents than the incremental patented inventions).

Therefore, at the invention level of analysis, the more *an invention* relies on scientific knowledge, the more pioneering it can be, and thus, probably, the higher its impact on subsequent inventions. But since, in this study, the unit of observation is the firm (and not the patented invention) and since, all the variables will be constructed on a yearly basis for each firm (firm-characteristics that are attributed to a given firm in a given year – see subsection *3.3* for details), we have to develop the hypotheses at the firm level of analysis. Consequently, at the firm level of analysis, we can argue that the more *a firm* builds on scientific knowledge to develop new technologies, the greater its impact on technological evolution.

The extant literature provides us with several studies that support this causality. For example, Miller et al. (2007) found a positive and significant relationship between the inventions that are linked to scientific research and their impact on subsequent inventions. Moreover, Gittelman and Kogut (2003) as well as Nagaoka (2007) demonstrated that incorporating scientific knowledge into inventions makes them more influential. Based on the above, we propose our first hypothesis:

Hypothesis 1a: *Ceteris paribus*, the more a firm relies on scientific knowledge to develop new inventions, the greater the firm's impact on technological evolution, as measured by the aggregate number of citations received by its patents.

Breadth of Impact: Theoretical knowledge is by definition more general than practical (i.e. technical) one, since practical knowledge can be generally considered as a result of the application of theoretical ideas to certain technological areas. The "descent" of an idea from a theoretical situation to a practical one, makes this idea less general. That is to say, a theoretical idea has a wider breadth of applicability compared to the applied version of this idea in a certain technological field, in which it has become more specific.

Moreover, the scientific knowledge within patented inventions (i.e. knowledge deriving from scientific journals) in comparison to the technical knowledge within patented inventions (i.e. knowledge deriving from prior patented inventions) is considered as more theoretical, and, therefore, this type of knowledge is expected to have a broader impact on technological evolution.

Consequently, inventions that rely heavily on scientific knowledge (i.e. on knowledge from scientific journals) are anticipated to be influential beyond their focal technological domain, since they incorporate many pieces of theoretical knowledge. Thus, for the above-mentioned reason, we can form, at the firm level of analysis, the following hypothesis:

Hypothesis 1b: *Ceteris paribus*, the more a firm relies on scientific knowledge to develop new inventions, the wider the breadth of the firm's impact on technological evolution.

2.5.2 Hypothesis 2 - Technological Collaboration

Magnitude of Impact: Technological collaboration is defined here as the number of partners (if any) with whom the focal firm cooperates with in order to develop and patent new technological inventions.

In the context of the increasing complexity of new technological projects and the decreasing availability of time for their accomplishment, firms view technological collaboration as a necessary practice for being at the cutting edge of technological achievements. The need for complementary resources and knowledge has led firms to seek for partners that can contribute to the generation, the development and the exploitation of new ideas (Chesbrough 2003). Hagedoorn (2002) demonstrated this trend clearly, by presenting the forceful growth of R&D partnerships during the last three decades.

Generally, the success of technological collaborations in creating new knowledge rests on the interplay between two antithetic forces that result from partners' diversity: *creativity* on the one hand and *transaction costs* on the other. Creativity here refers to the potential of arriving at novel combinations as a result of bringing different actors with diverse but complementary ideas, experiences, resources and capabilities, etc. Transaction costs, on the other hand, accrue precisely because of the diversity of actors: the more different the actors, the more difficult for them to communicate and coordinate effectively because of their diverse backgrounds, motives, or perhaps because of incongruent hidden agendas.

Nooteboom et al. (2007) stressed in depth this issue by examining the relation between cognitive distance and innovation performance. Cognitive distance refers to the differences between organizations in their "interpretation systems" (Weick 1979, 1995), their "systems of shared meanings" (Smircich 1983) or organizational "focus" (Nooteboom 2000), established by means of shared fundamental categories of perception, interpretation and evaluation inculcated by organizational culture. Nooteboom (1992, 1999) proposed that there is an inverted-U shaped relationship between cognitive distance and innovation performance. In first instance, as cognitive distance increases, it has a positive effect on learning by interaction. When people with different knowledge and perspectives interact, they stimulate and help each other to stretch their knowledge for the purpose of bridging and connecting diverse knowledge. However, at a certain point cognitive distance becomes so large as to preclude sufficient mutual understanding needed to utilize those opportunities. Of course, a certain mutual understanding is needed for collaboration, and familiarity certainly breeds trust (Gulati 1995), which facilitates successful collaboration. However, too much familiarity may take out the innovative steam from collaboration. The challenge then is to find partners at sufficient cognitive distance to tell something new, but not so distant as to preclude mutual understanding.

A small cognitive distance allows greater comprehensibility, but constraints novelty. Conversely, a large cognitive distance yields to novel knowledge, but limits comprehensibility. Partners sharing similar cognition have similar perceptions, interpretations and evaluations. Therefore, they understand each other actions and expressions. This implies that they can tell each other something new (although related to the partner's cognitive framework) and still communicate smoothly on the grounds of their common background. Therefore, a certain degree of cognitive distance is needed since it ensures that firms can connect their cognitive frameworks and being innovative as well as they can easily communicate. Conversely, if firms share exactly the same cognition, there will be a reduction of the innovative potential as they will think alike, while, if their cognitive distance is greater, there will be a difficulty in communication.

The above-mentioned interplay between creativity and transaction costs will be also present in the "collaboration for the development of new patented inventions", since this type of collaboration can be seen as a specific type of the more general "technological collaboration". When the number of partners with whom the focal firm cooperates with in order to develop and patent new technological inventions increases, the creativity of the team increases as well. This means that the different partners characterized by diverse ideas, experiences, resources and capabilities will have the potential to solve more complex and demanding technological problems, and consequently, to generate more groundbreaking and important inventions (which, in turn, will probably have a greater impact on technological evolution). However, when the number of partners increases beyond a certain point, the cognitive distance will become large and the transaction costs will be so high that would probably have a negative effect on the "quality" of the patents (and hence –ceteris paribus – on the number of citations they receive).

Therefore, we expect that as the number of the co-assignees (i.e. the number of partners that co-developed the patent) increases, the chances for impactful inventions will increase as well, but up to a point. After that point, the more co-assignees will affect negatively the firm's capability to develop influential inventions.

The empirical evidence on the relation between these specific variables (i.e. the number of co-assignees and the patent's citations received) are mixed and quite rare. Sapsalis et al. (2006) found that the number of organizations that are collaborating to produce a patent has a highly significant and positive impact on the patent's influence on future patents. However, Srivastava and Gnyawali (2011) showed that the number of patents' co-assignees is not associated to the number of citations that these patents receive.

All the above considerations lead us to the following hypothesis:

Hypothesis 2a: *Ceteris paribus*, a firm's impact on technological evolution, as measured by the aggregate number of citations received by its patents, is related to the number of the knowledge generating organizations with whom a firm collaborates to develop new inventions in a curvilinear (inverted-U shaped) manner.

Breadth of Impact: Concerning the effect of technological collaboration on the breadth of impact, it is logical to assume that, *ceteris paribus*, the inventions that are the result of the joint endeavor of two or more firms are built on more knowledge components than the inventions that resulting from the R&D activities of a single firm. The basis of this assumption is that in the collaborative R&D projects one of the basic challenges is to successfully merge the technological specializations in different areas of expertise of each participant, in order to be developed a technologically advanced invention. It is this variety of the areas of expertise of each participant that leads to the assumption that the knowledge components of the inventions that result from collaborative R&D projects are generally more than the knowledge components of the inventions that result from the R&D department of a single firm.

Moreover, we expect that the breadth of an invention's impact is positively related to the number of the knowledge components that are embedded within it. The logic of this proposition is the following: *ceteris paribus*, the higher the number of the different knowledge components that are embedded within an invention, the higher the number of the different technological classes to which these components belong, and, thus, the more the technological classes on which this invention may have an impact. In other words, the more diverse the knowledge upon which an invention is built, the more the technological fields to which this invention refers and might exert an influence on.

Moreover, an invention, just because of belonging in more technological areas increases the likelihood of affecting an explorative research project in a different technological field (since this explorative research project can “meet” the diversified invention in multiple technological fields), as compared to an invention that belong to a single technological area.

Therefore, the inventions that are the outcome of collaborative R&D projects are expected to exert a broader influence on technological evolution compared to single-firm inventions, since these inventions are composed of more numerous knowledge components which, in turn, probably belong to more technological classes; and, so, they can potentially affect the future patents of these particular technological classes. In addition, their presence in multiple technological classes increases their probability to be cited by an explorative invention from a different technological class in comparison with the inventions that belong to a single technological class. Consequently, these arguments lead to the following hypothesis:

Hypothesis 2b: *Ceteris paribus*, the more the knowledge generating organizations with whom a firm collaborates to develop new inventions, the wider the breadth of the firm’s impact on technological evolution.

2.5.3 Hypothesis 3 - Technological Diversity

Magnitude of Impact: By *Technological Diversity* we refer to the breadth of the technical knowledge within a firm (George et al. 2008). Many firms, in their effort to produce new products and services, are required to operate using a broad range of knowledge, based on the expectation that diversified technologies will lead to economies of scope and to technological synergies (Miller et al. 2007).

Firms with extensive technological diversity are better able to absorb external knowledge (Operti and Carnabuci 2014), to develop new combinations of different technologies by transferring knowledge components from one technological domain to another and, therefore, to generate more groundbreaking inventions. Furthermore, just because an invention belongs to a broader range of technologies can make it more influential since it can affect a larger number of future inventions (Sorenson and Fleming 2004) or, in other words, it actually increases the pool of inventions from which this invention can "attract" citations.

Consequently, the broader the range of technologies upon which an invention is based, the greater the invention's impact, and hence, at the firm level of analysis, the more technologically diversified a firm is, the greater the firm's influence on technological evolution.

In general, the existing literature supports the positive association between firm's technological diversification and its impact on technological change. George et al. (2008) found a positive and significant relationship between firm's breadth of technological capabilities, measured as the sum of the technical classes in which a firm had applied for patents, and firm's technological impact, measured as the number of citations received by all the patents filed by a firm in a year. Moreover, Fleming (2001) and Sorenson and Fleming (2004), controlling for the number of patent classes into which the patent falls (i.e. the number of technological domains to which a patent belongs), found a positive and significant correlation with the patent's impact on future inventions. All these suggest the following hypothesis:

Hypothesis 3a: *Ceteris paribus*, the more technologically diversified a firm is, the greater the firm's impact on technological evolution, as measured by the aggregate number of citations received by its patents.

Breadth of Impact: All else being equal, we expect the technological diversification of a firm to have a positive effect on the breadth of the firm's impact on technological change. In other words, we predict that the more the technological domains on which the inventions of a firm are based on, the more wide the breadth of the firm's impact.

The rationale behind this prediction is similar with the logic of the *hypothesis 2b*: firms with high technological diversification conduct broader search and develop inventions that belong to many technological domains, on which these inventions can exert an influence (Banerjee and Cole 2010). Moreover, the technologically diversified inventions just because of being present in many technological areas their chances of being cited by future explorative inventions from different technological areas (that seek for useful knowledge outside of their boundaries) are greater compared to the inventions that belong to a few technological domains, since those explorative inventions can "meet" the diversified inventions in more numerous technological classes. Consequently, it is natural to expect the breadth of an invention's impact to be commensurate to the breadth of the technological domains on which it applies.

With regard to the empirical evidence concerning this relationship, Argyres and Silverman (2004) supported this expectation. Measuring the breadth of technological impact as a Herfindahl Index and the technological diversification as a Concentric Index

[10], they demonstrated a positive influence of firm's technological diversification on the broadness of its technological impact. Moreover, Banerjee and Cole (2010) found that the technological diversification, computed as a Herfindahl Index based on the *citations made* - in our study we use the term *Originality* for this measure -, exerts a positive and significant impact on breadth of impact, measured as a Herfindahl Index based on the *citations received* (*Herfindahl Type Breadth*). Therefore, on the basis of the above, we formulate the following hypothesis:

Hypothesis 3b: *Ceteris paribus*, the more technologically diversified a firm is, the wider the breadth of the firm's impact on technological evolution.

2.5.4 Hypothesis 4 - New-to-Firm Technological Domains

Magnitude of Impact: This variable measures the degree to which a firm enters new-to-firm technological domains. Within a rapidly evolving technological context, firms, in order to avoid becoming technologically (and thus competitively) obsolete, are constantly searching for new-to-firm technologies that will lead them to successful innovations and, therefore, to competitive advantage.

The coupling of the knowledge which comes from a new-to-firm technological domain with the firm's existing stock of knowledge enables the R&D unit to experiment with novel combinations of knowledge that could result in breakthrough inventions. Ahuja and Lampert (2001) found some support for this prediction. In particular, their findings indicated that experimenting with new-to-firm technologies increases the likelihood of breakthrough inventions up to a point and then decrease (an inverted U-

[10] The Concentric Index of Diversification has been developed by Caves et al. (1980) and its differences from the Herfindahl Index are that it doesn't employ the squares of the technological classes and it takes account also of the "distance" between the different technological classes. See Caves et al. (1980, p. 199) for more details.

shaped effect). According to them, the positive effect happens because the exposure to different technological approaches enhances the repertoire of solutions that an organization can use to face any new problem and deepens its technological insights so as to encompass both old and new knowledge. On the other hand, the negative effect happens when a firm enters new-to-firm technologies to such an extent that can become a source of confusion and information overload.

In line with this reasoning, it is our view that whenever an inventor attempts to create new knowledge from the synthesis between new-to-firm and existing technologies, he/she has to be able to make correct and timely decisions with regard to whether this synthesis can lead to a worthy invention or it must be abandoned. In this phase, as long as the inventor deals with small or moderate quantities of new knowledge components, he will probably achieve to decide successfully and on time whether the new synthesis worth implementing. However, when the inventor deals with new-to-firm technologies in excess quantities, he/she will likely face difficulties in clarifying the new synthesis and in estimating its potential benefits and, in turn, this will probably affect negatively the R&D performance (more costly and timely developments ending with results that are below expectations) and will lead to inventions of lower quality and, eventually, to inventions of lower impact.

Furthermore, during the phase of the implementation and the development of an invention that is based on both existing and on new-to-firm technologies, when the exposure to new technologies does not exceed a certain level, the inventors would manage to combine knowledge components from the existing technologies with knowledge components from the new technologies in such a way that will probably lead to inventions that will be characterized by pioneering technological elements, and therefore, will be expected to exert a greater influence on future inventions (Ahuja and

Lampert 2001). But, the excessive exploration of new technologies could eventually proved to be harmful, since the simultaneous process of numerous knowledge components can cause confusion to the inventor and can lead only to a superficial understanding, a situation that will definitely affect negatively the quality of the invention and thus its impact.

Therefore, we expect that as the experimentation with new-to-firm technologies increases, the chances for impactful inventions will increase as well, but up to a point. After that point, the more experimentation with new-to-firm technologies will affect negatively the firm's capability to develop influential inventions. Based on the above, we formulate the following hypothesis:

Hypothesis 4a: A firm's impact on technological evolution, as measured by the aggregate number of citations received by its patents, is related to the firm's experimentation with new-to-firm technologies in a curvilinear (inverted-U shaped) manner.

Breadth of Impact: We anticipate that, *ceteris paribus*, firms that experimenting with new-to-firm technological domains, compared to those that do not, are expected to exert more broad influence on technological change.

Before we proceed to our argument, we remind readers that this variable measures the degree to which a firm enters new-to-firm technological domains and, as we will show in subsection *3.3.2*, we consider that a firm is entering a new technology domain when it applies for a patent in an technological class in which it had not applied in the previous 4 years, on the basis of assumption that technological classes in which a firm has not been active in the last 4 years are considered to be new-to-firm.

Therefore, a firm, by entering a technological domain in which it has never been active, increases the total number of technological domains to which its inventions belong. And, as we showed earlier (i.e. Hypothesis 3b), all else being equal, as the technological domains of the inventions of a firm increase, the technological domains that these inventions can have an impact also increase, and additionally, the likelihood to be cited by a patent from a different technological area increases as well.

Consequently, whenever a firm enters a new technological class and develops new inventions within that class, the number of the set of technological classes which are embedded within firms' inventions increases and, its technological effect becomes broader.

Banerjee and Cole (2010) have supported this claim by showing that the entry into a new technological domain has a positive and significant impact on the breadth of impact, measured as a Herfindahl index. Banerjee and Cole (2010) considered that a firm moves into a new technological domain when it develops a patented invention in a technology domain which is different from all past patented technology domains of the inventing firm.

Based on the above, we can formulate the following hypothesis:

Hypothesis 4b: *Ceteris paribus*, the more a firm experiments with new-to-firm technologies, the wider the breadth of the firm's impact on technological evolution.

2.5.5 Hypothesis 5 - Unexplored Technological Areas

Magnitude of Impact: This variable captures the degree to which a firm experiments with technological spaces that have not been explored before in a large extent, or, in other words, with technological spaces in which the prior research endeavors are rather limited.

Regarding the causal relation between the experimentation with unexplored technological areas and the technological impact, we argue the following: On the one hand, when developing technologies that have been previously thoroughly explored and in which there exists a plethora of technological antecedents, the use of knowledge elements that are known to have succeeded in the past provides to the inventor a base of familiarity from which he/she can take a step forward and, also, some assurance that his/her efforts will not result in complete failure. On the other hand, by working on technologies that have limited technological antecedents, the inventor moves to a largely unexplored technological area in expectation of completely new solutions and revolutionary ideas. Although this movement does increase the likelihood of a complete failure, at the same time increases the likelihood of a groundbreaking invention which will probably have a large technological impact (Ahuja and Lampert 2001).

With regard to the existing empirical evidence, Ahuja and Lampert (2001) identified the extent to which a firm experiments with unexplored technological areas[11] as a critical determinant of the firm's ability to shape technological evolution. Following this line of inquiry, we offer the following hypothesis:

[11] Ahuja and Lampert (2001) refer to the experimentation with unexplored technological areas as "Pioneering Technologies". However, they measure it in a slightly different way. Ahuja and Lampert (2001) measure it as the number of patents that cite no other patent. We measure it as the average number of citations a firm's patents make to previous patents, following other studies such as Sapsalis et al. (2006) and Operti and Carnabuci (2014). The main advantage of the measure that we employ is that it distinguishes between patents that cite just a few prior patents from patents that cite a lot of prior patents and, so, it accounts for the *quantity* of the prior art citations. On the other hand, the main advantage of the measure of Ahuja and Lampert (2001) is that it can capture the *pure* "Pioneering Technologies" that have zero technological antecedents. See more in subsection *3.3.2*.

Hypothesis 5a: *Ceteris paribus*, the more a firm experiments with largely unexplored technological areas, the greater its impact on technological evolution, as measured by the aggregate number of citations received by its patents.

Breadth of Impact: We anticipate that the extent to which a firm experiments with largely unexplored technological areas will exert a positive influence on the breadth of technological impact.

The reason for the above expectation rests on the nature of the explorative research projects. Those projects can be considered as more risky than the exploitative research project, since their quest for useful knowledge elements beyond their traditional technological areas denotes a tendency for breakthrough and radicalness. At the time when the inventors that participate in an explorative research project have to assess whether an invention from a *different* and *unexplored* technological area could be useful or not for their needs they will face a dilemma. On the one hand, they might be cautious about building their research on technologies that have not been tested thoroughly in any technological field, but, on the other hand, the completely new solutions and the revolutionary ideas that are embedded within these pioneering inventions will make them to appear "attractive" and promising in the eyes of these inventors. And because of the "risky" nature of the explorative research projects, their inventors will generally choose to build on inventions that are located in unexplored technological areas rather than to build on inventions that are located in fully-explored technological areas.

Therefore, the inventions that belong to unexplored technological spaces will exert a broader influence on future inventions that the inventions that belong to technological spaces with numerous technological antecedents. Thus, on the basis of the above syllogism, we propose the following hypothesis:

Hypothesis 5b: *Ceteris paribus*, the more a firm experiments with unexplored technological areas, the wider the breadth of the firm's impact on technological evolution.

2.5.6 Hypothesis 6 - Recency of Technological Inputs

Magnitude of Impact: Recency of Technological Inputs measures the recency (or the maturity) of the technologies that a firm incorporates into its inventions (Sørensen and Stuart 2000).

The work of Heeley and Jacobson (2008) stressed that for older technological inputs there is less likelihood of new knowledge recombination, because, the more mature a technology, the less useful the combinations of its elements that have not been *already* explored. This lack of potential for useful novel combinations results naturally in lower potential for influential inventions. In the same vein, Ahuja and Lampert (2001), based on the technology life cycle approach, argued that emerging technologies, being in the early stages of their life cycle, are characterized by high uncertainty and by a series of significant unsolved problems. But it's exactly these fundamental technological problems that can lead to solutions that may be of a path-breaking and impactful character. On the contrary, as a technology matures, the fundamental problems that remain to be solved decrease. Therefore, the possibility of creating groundbreaking and influential inventions is higher in emerging technologies than in mature ones.

Previous studies have generally demonstrated a positive link between the recency of the technologies that a firm incorporates into its inventions and the firm's impact on technological evolution. For example, Cattani (2005) found evidence that firms that build on older technologies are more likely to generate innovations with less impact (Conversely, firms that build on more recent technologies are more likely to generate influential innovations.) Moreover, Miller et al. (2007) and Joshi and Nerkar (2011) showed a negative relationship between the maturity of the inventions' technological inputs and the impact of these inventions on future inventions. The same results also appeared in the study of Rosenkopf and Nerkar (2001), with the difference that the impact of the focal inventions on future inventions concerned only the future inventions that belong to the same technological domain as the focal inventions. Consequently, on the basis of the above, the next hypothesis is set forth:

Hypothesis 6a: *Ceteris paribus*, the more recent the technologies that a firm incorporates into its inventions (as reflected in the backward citations of its patents), the greater the firm's impact on technological evolution, as measured by the aggregate number of citations received by its patents.

Breadth of Impact: We expect the recency of the inventions' technological inputs to be associated with the breadth of their impact. In particular, we predict that the more recent the technologies that are embedded within an invention the more broad the invention's technological impact.

To support this argument, we rely on the following rationale: when the inventors of an explorative research project are searching for useful knowledge components beyond their technological areas, the mature technologies may appear to them less "attractive" compared to recent technologies. Probably, this happens because the inventors assume that the more mature a technology, the less useful the combinations of its elements that have not been *already* explored and the less the likelihood a piece of knowledge from the mature technology to be successfully combined with their own explorative research effort. Thus, the knowledge components of an invention that relies fully or heavily on mature technologies (as reflected in the prior art on which it is based) may appear to them less promising and as something that is not worth the effort.

Therefore, inventions that are based heavily on mature technologies are not expected to attract explorative research projects from different technological fields. Conversely, inventions that rely largely on recent technologies, even if they may transmit signals of uncertainty, their potentiality and their fertile ground for groundbreaking and impactful technological developments cause them to appear attractive to explorative research projects from different technological areas.

Therefore, we expect the *Recency of Technological Inputs* to be positively related to the breadth of impact, even if Banerjee and Cole (2010), who used the *Recency of Technological Inputs* in their model as a control variable, reported a negative relationship between this variable and the Herfindahl type breadth of impact.

Thus, based on the above argumentation, we propose our next hypothesis:

Hypothesis 6b: *Ceteris paribus*, the more a firm incorporates recent technologies into its inventions, the wider the breadth of the firm's impact on technological evolution.

2.5.7 Hypothesis 7 - Recent Technological Activity

Magnitude of Impact: Recent Technological Activity measures the size of the knowledge base and knowledge stock upon which a firm builds its new technological inventions and it can indicate its technological competence and its technological expertise (Phene et al. 2006, Hoetker and Agarwal 2007, Operti and Carnabuci 2014*)*. This variable - which has been labeled with a variety of names such as patent stock (Phene et al. 2006), technological capital (Vanhaverbeke et al. 2009), prior capabilities (Nerkar 2003), and knowledge base size (Operti and Carnabuci 2014) - is measured on the basis of the number of the inventions that have been developed by a firm in the five years previous to the year of observation (Vanhaverbeke et al. 2009).

Because of the cumulative character of technology, the current technological activity of a firm is dependent on its previous level of technological activity (Vanhaverbeke et al. 2009). That is to say, the number of technological inventions that a firm develops in a given year is positively related to the number of the inventions that the firm produced in the prior years. And since patents citations are strongly related to the number of patents (the more the patents of a firm, the more the citations it receives), we expect *Recent Technological Activity* to exert a positive influence on technological impact. However, in our model, this effect cannot be observed because we will control for the number of patents. That is to say, the control variable *Number of Patents* will absorb this specific kind of impact that *Recent Technological Activity* has on the *Number of Patent Citations* (i.e. the effect on patent citations by affecting the number of patents).

Nevertheless, we anticipate *Recent Technological Activity* to have a positive impact on patent citations not by affecting the number of patents, but by affecting the firm's technological reputation and attractiveness to other firms. According to Hoetker and Agarwal (2007, p. 454), "technologies developed by firms perceived as highly technologically active might draw disproportionate attention from other firms". The reputation of the technological capabilities of a firm is logical to be strongly and positively associated with the results of the firm's R&D endeavors — namely the number of the inventions that have been developed by the firm in the previous years. Therefore, the inventions produced by high-reputation firms (reputation with regard to their technological capabilities) are expected to be more attractive and more trustworthy in the eyes of the inventors who are developing new technologies as compared to the inventions produced by low-reputation firms.

The claim that inventions which belong to more technologically active firms are cited more heavily has been empirically supported by Hoetker and Agarwal (2007), who found that recent technological activity (three years prior to the year of observation) is positively related to the number of patent citations received by the patents of a firm and by Srivastava and Gnyawali (2011), who showed that recent technological activity (five years prior to the year of observation) is positively associated with highly-cited patents.[12].

Consequently, on the basis of the above considerations, we propose the following hypothesis:

Hypothesis 7a: *Ceteris paribus*, the higher the recent technological activity of a firm, the greater the firm's impact on technological evolution, as measured by the aggregate number of citations received by its patents.

Breadth of Impact: As we noted previously, in *hypothesis 7a*, although *Recent Technological Activity* can affect the number of patent citations due to the cumulative character of technology (i.e. the higher the *recent* technological activity, the higher the *current* technological activity, and, thus the greater the technological impact), this effect will not be revealed in our model since we will control for the current technological activity (number of patents). Nonetheless, as we argued earlier, we anticipate *Recent Technological Activity* to have an impact on the other firms' perception of the technological capabilities of a given firm, and consequently, to have an impact on the extent to which the other firms trust the focal firm's new technological inventions and decide to build upon them.

[12] However, we have to mention that none of the two studies control for the number of patents. So, the impact of *Recent Technological Activity* on the *number of patent citations* might be due to the effect of *Recent Technological Activity* on the number of patents produced and not due to the effect on technological reputation and attractiveness.

We expect that this effect will be even more pronounced when the other firms belong to different technological domains. In this case, the inventors who seek for useful knowledge beyond their own technological areas will rely more on reputations with regard to the technological competencies of a firm and less on specific technical criteria (since they are not expert in those fields), as compared to the inventors who belong to the same technological area. Therefore, we anticipate firms with high *Recent Technological Activity* to exert a broader impact on technological evolution, by influencing the firms' technological reputation especially in the eyes of the inventors from different technological areas. Building on the above syllogism, we propose our next hypothesis:

Hypothesis 7b: *Ceteris paribus*, the higher the recent technological activity of a firm, the wider the breadth of the firm's impact on technological evolution.

2.5.8 Hypothesis 8 - Internal Focus

Magnitude of Impact: Internal Focus refers to the degree to which a firm builds upon and elaborates upon its prior research endeavors (Sørensen and Stuart 2000). Firms with strong *Internal Focus* are firms which strategically choose to build upon technologies that have been developed by themselves.

Concerning the relationship of *Internal Focus* to technological impact, Rosenkopf and Nerkar (2001) stressed that the inwardly focused exploration for new inventions even if it is one of the hallmarks of core competence (i.e. a firm develops its core competence by building on the knowledge of its previous work), at the same time it can lead to competency traps and to core rigidities. By building upon their own technologies and by neglecting to watch carefully and to understand in depth the developments of other firms in their own technological domain or in relatively close technological domain, firms are in danger of becoming technologically isolated. Assuming that technology evolves rapidly (especially in the high-tech industries) and new competitive technologies emerge continuously, the firms that are strongly stuck to their own technological trajectories without integrate external developments run the danger of a myopic behavior towards the new and the more promising technological trajectories that are being arose (this will also be the case in the long-run, even if the focal firms' technological trajectories were very important in the beginning of their life-cycle). This phenomenon can lead to a situation where the technologically isolated firms will develop inventions within technological areas which have been abandoned to a large extent by other firms. Thus, the inventions of such firms will probably affect only a few future inventions. Therefore, we can argue that the inwardly focused exploration of a firm will negatively affect the firm's ability to shape technological evolution.

With respect to the empirical findings on this relationship, Rosenkopf and Nerkar (2001) provided evidence for the negative effect of internal focus on the ability of a firm to influence technological change, by showing that the exploration within organizational boundaries has less impact on subsequent technological evolution than exploration that spans organizational boundaries. Therefore, on the basis of these arguments, we propose *Hypothesis 8a*:

Hypothesis 8a: Ceteris paribus, the more a firm builds on technologies that have been develop by it, the less the firm's impact on technological evolution, as measured by the aggregate number of citations received by its patents.

Breadth of Impact: Previously, we defined *Internal Focus* as the degree to which a firm builds upon prior technologies of its own to create new inventions. As we detail below (see subsection *3.3.2*), *Internal Focus* is measured by the ratio of the citations made to its own patented inventions to its total citations made, that is to say, by self-citations (Hoetker and Agarwal 2007).

When developing *Hypothesis 8a*, we argued that firms which build extensively upon their own technologies run the risk of being stuck on a technological trajectory which is gradually abandoned by other competitors (which have, apparently, followed new technological trajectories). Thus the more inwardly focused the inventions of a firm are, the greater the danger of becoming technologically isolated, which may result in a situation where there will be no firm to be interested in, or to be influenced by the inventions of the focal firm.

On the effect of internal focus on the breadth of technological impact, we argue that the aforementioned phenomenon will be even more intense when we are interested in the firms from different (and "distant") technological domains. If firms that operate in the same technological field with a given firm raise objections to follow the firm's technological developments that are becoming gradually technologically obsolete, then the firms from different technological areas will raise even stronger objections. That is to say, they will be even more unwilling to follow technologies that are technically unfamiliar with them and at the same time, these technologies are abandoned by firms from the same technological area (firms that are technically familiar with these technologies). Thus, it is logical to hypothesize that the more extensively a firm draws upon its own technological achievements, the more restricted the breadth of its technological impact.

With regard to this specific relationship, Rosenkopf and Nerkar (2001) provided some empirical evidence in support of our claim. In particular, by using the same measurement (i.e. self-citations) to compute *Exploration within Organizational*

Boundaries (they used this name instead of *Internal Focus*), they found that the more the *Exploration within Organizational Boundaries*, the less the impact on subsequent technological evolution beyond the focal domain. On the basis of their finding, the authors concluded that "apparently, firms that focus inward on their core competencies run the risk of developing innovations that wind up being peripheral to the aggregate path of technological development" (Rosenkopf and Nerkar 2001, p. 303).

Consequently, building upon the above, we propose *Hypothesis 9b*:

Hypothesis 8b: Ceteris paribus, the more a firm builds on technologies that have been develop by it, the more limited the breadth of the firm's impact on technological evolution.

2.5.9 Hypothesis 9 - R&D Investment

Magnitude of Impact: R&D Investment is a measure of the amount of resources that a firm decides to allocate to the technology search process (Yayavaram and Ahuja 2008, Operti and Carnabuci 2014). A sizable literature has documented the relation between a firm's R&D investment and its patenting activity (Cattani 2005), expressed, mainly, in terms of patents count (Ahuja 2000) or citation-weighted patent counts (Operti and Carnabuci 2014).

R&D Investment is likely to be a significant determinant of the number of patent citations, since the greater the investment of a firm in R&D, the more frequent the firm's patenting activity (Ahuja 2000), and, thus, the more the citations received by its patents. However, in our case, we will control for the number of patents that a firm develops, and, consequently, the impact of *R&D Investment* on patent citations via patent counts will be eliminated. Nevertheless, the *R&D Investment* can have a direct impact on patent citations even if the patent counts are controlled for. The rationale for this argument is that high R&D investment reveals an explorative inventive strategy which aims more at radical and breakthrough inventions rather than incremental ones. In other words, when two firms from the same technological field develop the same number of patents but the first firm has spent more in R&D than the second, then, it is logical to believe that the patents of the first firm will be on average more radical than the patents of the second firm. The logic of this assumption is that radical inventions, as a rule, require more expenses than incremental ones. Therefore, firms that invest more in R&D are expected to develop more radical and breakthrough inventions, *ceteris paribus*, and, thus, to exert a greater influence on the evolution of technology.

To date, the literature on this topic has presented evidence that support our claim. In particular, Cattani (2005) found R&D expenditure to be a significant predictor of the impact of a firm's patenting activity while Yayavaram and Ahuja (2008) showed that firms with higher R&D intensity tend to develop inventions of greater impact[13]. Based on the above, we propose the following hypothesis:

Hypothesis 9a: *Ceteris paribus*, the more a firm invests in R&D, the greater the firm's impact on technological evolution, as measured by the aggregate number of citations received by its patents.

Breadth of Impact: In *Hypothesis 9a*, we stressed that R&D investment may have a positive impact on the number of patent citations but not via the quantity of patents produced (since we control for the number of patents). This impact may occur via the quality of patents. That is to say, all else being equal and controlling for the number of

[13] Also, at this point, it is important to mention that both of these studies did not control for the number of patents.

patents, high R&D investment may denote a research strategy that focuses on the development of explorative and radical inventions.

In addition, radical inventions tend to evidence a broader technological impact compared to incremental inventions, whose impact can be characterized as predictable and fixed in terms of breadth of impact. We anticipate that the attention of inventors who seek for pieces of knowledge beyond their technological area will be attracted by the radicalness of the breakthrough inventions. And, if these novel pieces of knowledge are considered by the inventors to be useful for their developments, they will try to adopt them and to combine them with their own research projects. Therefore, we expect firms with high R&D investment, by developing radical inventions, to exert a broader impact on technological evolution.

However, the only relevant empirical finding on this particular relationship does not support our argument. In particular, Valentini and Di Guardo (2012), presented that the R&D intensity, measured as R&D expenditures over sales, does not affect the Herfindahl type of breadth of impact (they also controlled for the number of patents).

Nevertheless, we still anticipate that high R&D investment will lead to a broader technological impact, and so, we propose our next hypothesis:

Hypothesis 9b: *Ceteris paribus*, the more a firm invests in R&D, the wider the breadth of the firm's impact on technological evolution.

2.5.10 Hypothesis 10 - Firm Size

Magnitude of Impact: In general, large firms are responsible for a disproportionate quantity of inventions as measured by the number of patents filed (Cattani 2005). Thus, it is logical to assume that they are exerting strong influence on technological evolution by receiving a large amount of patent citations. However, since in our model we will control for the number of patents, this particular effect of firm size on the number of patent citations will not be revealed. Nonetheless, we expect firm size to affect patent citations in others ways that can be either negative or positive.

On the one hand, with regard to the negative effect of firm size on patent citations, Sørensen and Stuart (2000) have stressed that the bureaucracy and the lack of entrepreneurial culture that often characterize large organizations may affect negatively the degree to which they can develop breakthrough (and thus influential) inventions. Moreover, it might be the inertia of the large firms that can make their knowledge base rigid and inflexible (Yayavaram and Ahuja 2008) and can inhibit explorative and groundbreaking research projects in favor of the specialization along existing technological trajectories (Vanhaverbeke et al. 2009). Thus, because of the reasons outlined above, we argue that a negative effect always inheres in the relation between firm size and patent citations.

On the other hand, with respect to the positive effects of firm size on patent citations, the reputation and the prestige of larger firms may affect positively the way in which other inventors assess their inventions. Inventors are often biased towards the inventions of larger and more prestigious firms as opposed to the inventions of smaller firms, and, thus, they are more willing to adopt and elaborate their technologies even if they are not absolutely sure about their technological superiority (Sørensen and Stuart 2000). Moreover, the economies of scope and scale of larger firms may have a positive effect on the quality of the R&D projects and they can lead to valuable and important inventions (Yayavaram and Ahuja 2008). Additionally, larger firms just being move "visible" than others firms may attract attention that translates into patent citations from other organizations (Argyres and Silverman 2004).

Moreover, another positive effect of firm size on technological impact could be that the size of a firm can be used as a negotiating tool in dealing with governments concerning regulations which are of major interest to the firm. More specifically, governments, under the threats of specific actions that can be taken by a large firm, such as mass layoffs, downsizing, and transferring business units to other countries (actions with a significant negative impact on employment and government revenues), may be pushed to proceed to favorable for the firm regulations. And, if these regulations regard technical standards then the technologies of this "privileged" firm will probably prevail over rivals' technologies.

Furthermore, larger firms because of their stronger financial power have an advantage over smaller firms in shaping technological standards by financing industry associations, promoting technological standards and financing lobbyism[14]. Especially, in technologically intensive industries, technological standards are of paramount importance, since those standards can determine and prescribe the future trajectory of a specific technology and force all the competitive technologies to adjust to the specific technological standards (see examples in the works of Rosenkopf and Tushman (1994) and Tushman and Murmann (1998)). Actually, the technological standards or whatever kind of institutionally established norm or requirement regarding technical systems can be viewed as a form of dominant design as we know it from the studies of Rosenkopf and Tushman (1994), Tushman and Murmann (1998) and Munir and Jones (2004). The "institutionalized" technology (e.g. the technology upon which a technological standard is based) appears to the inventor of the future patents always as more secure, more tested

[14] See Murmann (2013) for the importance of lobbyism for the creation of a munificent environment for the firms at the industry level of analysis.

and more trustworthy, consequently, the “institutionalized” technologies are expected to exert a greater influence on technological evolution.

As far as the empirical evidence of this particular relationship are concerned, prior research provides us with evidence for both positive and negative effects. On the one side, Yayavaram and Ahuja (2008), found firm size to be a significant positive predictor of technological impact. On the other side, Argyres and Silverman (2004), reported a negative and significant effect of firm size on the ability of a firm to influence technological evolution. However, it is important to note that none of these studies controlled for the number of patents.

In our case, considering the significance of the factors related to firm size that affect positively the technological impact contrary to those that affect it negatively, we expect positive impacts to be stronger than the negative and we anticipate this to be revealed in our models. Therefore, all the above considerations lead us to the following hypothesis:

Hypothesis 10a: *Ceteris paribus*, the larger a firm, the greater the firm’s impact on technological evolution, as measured by the aggregate number of citations received by its patents.

Breadth of Impact: When developing *Hypothesis 10a*, we argued that we expect larger firms to have a greater ability to shape technological evolution by producing patents that are more “institutionalized” and of higher reputation, prestige and quality. These qualitative characteristics can make an invention to be more “visible” and more attractive in the eyes of the inventors of the future patents.

We argue that this would also be the case with a specific category of inventors, the inventors from different technological domains. When a researcher in a distant technological area is focusing on a particular technology in order to discover a useful for his/her purposes piece of knowledge, the reputation that accompanies the inventions of large organizations may prove to be decisive for the inventors' choice. Or, it might be the technological superiority of the inventions of the large firms. Larger firms, by taking advantage of their economies of scope and scale, may have the potential to develop inventions of higher quality, as compared to smaller firms. Thus, on the basis of technological criteria, the inventors might prefer the large firms' patents to build upon.

Moreover, it could be the "better institutionalization" of the big firms' technologies that will determine the choice of the "distant" inventor. Large firms, by threatening with mass layoffs, downsizing, and transferring business, or by using their financial power to promote technological standards and to finance industry associations and lobbyism can put pressure on governments to regulate in a "friendly" manner. Thus, especially for the "distant inventors" who are unaware of the details of how certain technologies have been institutionalized in a foreign to them field (due to their technological superiority or due to their "social" power to prevail?), the degree of institutionalization of a technology can affect significantly their selection. Consequently, in the eyes of the inventors from different technological areas the "institutionalized" technologies will always outweigh the "non-institutionalized" technologies.

Extant empirical research provides some support for this assertion. More specifically, Valentini and Di Guardo (2012) showed that the firm size, measured as the log of assets, exert a positive and significant influence on the breadth of impact, computed as a Herfindahl index.

Therefore, taking into consideration all the above factors, we formulate our final hypothesis:

Hypothesis 10b: *Ceteris paribus,* the larger a firm, the wider the breadth of the firm's impact on technological evolution.

CHAPTER 3

RESEARCH METHODOLOGY

3 RESEARCH METHODOLOGY

3.1 Research Setting

To empirically test our hypotheses, we chose the pharmaceutical, the biotechnology and the chemical industries as our research setting. A number of reasons motivated the choice of these industries as the appropriate empirical context for the study. First, within these industries, there exists intense technological competition (De Carolis 2003, Gittelman and Kogut 2003) which, in turn, makes technology to evolve rapidly and the seven-years period of our study, from 2003 to 2009, (in subsection 3.2.3, we discuss more analytically the specific time period that is covered in our study), an adequate time period for the examination of technological evolution. Second, these industries are three of the most invention-intensive industries, and they are characterized by high R&D intensities (Powell et al. 1996, Rothaermel and Hess 2007). Third, within our selected industries the intellectual property protection is particularly strong, and thus, the link between patents and inventions is likely to be also strong (Ahuja 2000, Sørensen and Stuart 2000).

Moreover, these industries are heavily reliant on science (George et al. 2008), with the relationship between scientific knowledge and patented inventions to be especially important (Gittelman and Kogut 2003). In addition, technological collaboration is a significant feature of these industries (Ahuja 2000), while the technology-seeking alliance is a usual tactic (Anand et al. 2010). For all the above-mentioned reasons, we believe that the global industries of pharmaceuticals, biotechnology and chemicals provide a natural setting to test our hypotheses regarding those characteristics that determine a firm's impact on technological evolution and the breadth of its technological impact.

3.2 Sample and Data

We focused on firms that invest systematically in R&D from the three industries discussed above. This means that our sample is not representative, since it consists only of large firms with high R&D expenditure, and the consequence of this is that all the statistical inferences will regard mainly this kind of firms (large firms with high R&D) and not any firm that belongs to pharmaceutical, biotechnology, or chemical industry.

To test our hypotheses, we drew on the *EU Industrial R&D Investment Scoreboard*, as well as on the *Derwent Innovation Index Database*, for the period from 2003 to 2009.

3.2.1 EU Industrial R&D Investment Scoreboard

The EU Industrial R&D Investment Scoreboard *(henceforth*, the *Scoreboard)* provides economic and financial data of the top *corporate* R&D investors from all over the world. It is part of the European Commission's monitoring activities to improve the understanding of trends in R&D investment by the private sector and the factors affecting it. The *Scoreboard* is published annually in order to provide a reliable, up-to-date benchmarking tool for comparisons between companies, sectors, and geographical areas, as well as to monitor and analyse emerging investment trends and patterns. Apart, from the data on *R&D Investment*, the EU Industrial R&D Investment *Scoreboard* provides information for each company about its *Operational Profit*, its *Net Sales*, its *Country of Origin*, its *Industry*, its *Number of Employees*, its *Operational Profit*, its *Market Capitalization*, and about its *Capital Expenditure*. The companies listed in *Scoreboard* account for more than 90% of worldwide business enterprise expenditure on R&D.

The data for the *Scoreboard* are taken from companies' publicly available audited accounts. Companies are categorized in industry sectors according to the NACE Rev. 2 (Nomenclature statistique des activités économiques dans la Communauté européenne) and the ICB (Industry Classification Benchmark). In order to maximise completeness and avoid double counting, the consolidated group accounts of the ultimate parent company are used and the companies which are subsidiaries of any other company are not listed separately. Thus, for firms that are diversified in various industries, their R&D expenditure concerns any R&D expense regardless of the industry.

The first *Scoreboard* was published in 2004 and concerns the financial year 2003. We created a sample of 139 firms from biotechnology (29 firms), pharmaceutical (46 firms) and chemical (64 firms) industries for the period from 2003 to 2009. From these industries, we selected all firms that are present every year in the *Scoreboard* so as to have a balanced panel (the same firms are present in every year of observation). As we will show in subsection *3.4.1*, this characteristic of our sample (balanced panel data) will protect our regression models from a certain type of statistical bias (Wooldridge, 2002). However, it will obviously lead to fewer observations in our sample and, more important, the inference that will be drawn from of our results will concern only a certain type of firms: firms with consistently high R&D expenditure.

Our sample consists mainly of large firms with high R&D investments since the average firm size is 16264 employees, the average annual R&D investment is 527 millions of euro and the average R&D investment relative to sales is 9% (based on the year 2009). Moreover, the patent activity of our sample's firms is relatively intense since in year 2009 each firm applied for 126 new technological patents on average. Most of the firms of our sample are based in Europe (75 firms) and in particular in Germany (18 firms), UK (15 firms) and France (8 firms). Moreover, 31 firms are based in USA while 32 in Japan. *Appendix A* presents the firms in our sample and their countries of origin, for each industry.

Continuing with some descriptive statistics for each of industry (*Table 3.1*), it is worth mentioning that biotechnology firms even if they are in average the smallest, they spend the highest percent of their sales in R&D. Moreover, it is interesting to note that chemical firms develop new technological patents spending less, as compared to biotechnology or pharmaceutical firms.

Table 3.1 Descriptive Statistics for each Industry based on year 2009

Industry	Number of firms	Average Number of Employees	Average R&D Investment	Average R&D to Sales	Average Patents Applied
Biotechnology	29	2505.28	222.60	20.21%	15.41
Pharmaceuticals	46	23007.93	1124.65	14.85%	59.67
Chemicals	64	17652.42	235.52	3.73%	224.28

3.2.2 Derwent Innovation Index Database

Our source for the patent data is the Derwent Innovation Index Database. The Derwent Innovation Index (*DII*) is a database of international patent information. Over 14.3 million basic inventions are covered in *DII*, with coverage dating back to

1963, from 40 worldwide patent-issuing authorities (Gittelman and Kogut 2003, Alencar et al. 2007, Lettl et al. 2009).

The main advantage of *DII* is that it gives information on the full patent family, comprising all the different patents issued in different jurisdictions around the world on a given invention and all the different patents that are considered as extensions (or improvement) of this invention (Gittelman and Kogut 2003). In other words, the patent family is a set of patents that are related to the same invention. And because, in this study, by examining the impact of an invention, we are interested in the invention considered as novel technical ideas and not in the invention considered as a legal document that grants exclusive use and rights to the inventor or assignee (i.e. patent), the use of patent family eliminates the bias introduced by the counting of the same invention multiple times (Lettl et al. 2009).

For example, firm *A* developed in year *t* 50 new technological inventions. To protect these inventions, firm *A* issued or renewed *100* patents in various patent offices worldwide. On the contrary, firm *B* developed 50 new technological inventions but it issued or renewed *200* patents in different countries for these inventions. It could be the different geographical coverage that each firm choose for the protection of its inventions (the more the countries in which an invention is protected, the more the patents related to the invention) or it could be the nature of some specific inventions for which a series of improvements can be gradually made (and, these improvement need also to be protected) that can explain the difference in the number of patents. So, these two firms developed in year *t* the same number of novel technological ideas (50) even if in terms of patents documents differ greatly. Therefore, patent families can capture new technological inventions unbiased by number of patent documents.

The use of patent families determines the way that *DII* constructs its data. All the measures are based on patent families and not on single patents. For example, as we will show later, the number of citations is computed as the number of citations that the firm's patent families receive from subsequent patent families. So, if a patent cite two patent documents that belong to the same patent family, *DII* count it as one citation. If two patents that belong to the same patent family cite another patent, *DII* count them again as one citation. In this way, we can capture the effect of a given technological invention on subsequent technological inventions and to avoid capturing the effect of a given patent document on subsequent patents documents.[15]

Another useful characteristic of *DII* is the use of the assignee code, which is a standardized form of patent assignee (Alencar et al. 2007). By using this code, all the subsidiaries and the related holdings of a company (regardless of the industry to which they belong) are sharing the same code with the parent company. This feature significantly simplifies our search for the patent families of each firm that is included in the sample[16]. Moreover, this feature of *DII* is in accordance with the way that our first data source, the EU Industrial R&D Investment Scoreboard, measures the R&D investment of each firm (the R&D investment of each subsidiary is added to the R&D investment of the parent company).

For each patent family, (i.e. for each invention) *DII* provides a rich set of data upon which we built our patent-based variables. We will proceed with a description of

[15] Although, the use of patent families is very convenient for us that we are interested in technological inventions, it might be inconvenient for other researchers that are interested, for example, in individual inventors. If a patent family contains three different patents and the two of them are improvements of the first patent, then an inventor that has participated in all three patents will be equal to the inventor that has participated in only one patent. *DII* lists the names of all the inventors that have participated in all the patents of a certain patent family but it does not list the inventors per patent. So, the inference about inventors based on *DII* may be problematic.

[16] We gather all the patent families that a firm has applied for regardless of the patents' technological class. That is to say, if a patent family of a diversified firm has been categorized in a technological class that does not belong to the industry of the firm (e.g. a chemical firm applies for pharmaceutical patents) is still taken into account.

a typical Derwent record in order to let the reader understand what kind of information is included within it. Later on, in *Table 3.2*, we will show how each of our variable that has been constructed upon the data from *DII* is related with certain fields of a Derwent patent record. The titles (given by *DII*) and the descriptions of the main data that are contained within a Derwent patent record follow:

- *Title*: A concise descriptive English-language title written by *DII* to highlight the content and novelty of the invention disclosed in the patent specification.
- *Patent Number*: All the patent documents that belong to the same patent family. The Patent Number is a serial number assigned to each patent document by the patent-issuing authority. *DII* enrich this serial number by inputting the two-character country code of the publishing country, followed by the above-mentioned serial number, and the status code indicating the document type or the publication stage.
- *Inventors*: The list of the names of the inventors that have participated in at least one of the patents of focal patent family.
- *Patent Assignee Names and Codes*: The corporate body to whom all or limited rights of the patent are legally transferred, along with a unique four-letter code that is assigned by *DII* in order to unify ambiguous assignee names. By using this four-letter code, we search for all the patent families that the firms of our sample are participating in.
- *Derwent Primary Accession Number*: A unique identification number assigned by *DII* to the first patent in each patent family, and therefore to the database record created for that family. The format of each number is the year of publication and a six-digit serial number.

- *Patents Cited by Inventor / Examiner:* The list and the number of patents cited by the inventor or by the patent examiner.
- *Citing Patents*: The list and the number of patent family records whose members have cited members of the focal patent family.
- *Articles Cited by Inventor / Examiner:* The list and the number of articles (non-patent items) cited by the inventor or by the patent examiner.
- *Abstract*: The abstract is prepared by *DII* after reviewing the claims[17] and the disclosure of the patent. Written in English, the abstract is concise, accurate, and relevant, covering the widest scope of the invention as set out in its main claim.
- *International Patent Classification (IPC)*: An internationally recognized classification system that is controlled by the World Intellectual Property Organization (WIPO) and assigned to patent documents by the Patent Office that publishes the document. The IPC covers all technologies and is a useful system for searching patents with greater precisio. For example, the IPC of Biotechnology include areas such as transgenic vertebrates, invertebrates and plants; methods, processes and testing; bioinformatics; biological materials, etc. These are: 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).
- *Patents Details*: For each patent that is included within a given patent family the following details are provided: i) *Patent Number*: A serial

[17] Each patent contains a set of claims which are the list of the specific technological developments for which the patent assignee is claiming exclusive rights (Harhoff and Wagner 2006). Patent claims actually declare the specific novelties that are claimed to have been achieved by a particular patent family (Markman et al. 2004).

number assigned to each patent document by the patent-issuing authority. ii) *Patent Publication Date*: The date on which the patent document is made available to the public. iii) *Main IPC*: The classification number of the patent according to the hierarchical classification system produced by the World Intellectual Property Organization. iv) *Derwent Week*: Represents the week the data was entered in the DII database.

- *Application Details and Date*: The application number is the local filing number assigned to the patent document by the patent office. The application date or filing date is the date on which the application was filed with the patent office. The filing of a patent application is normally in the applicant's domestic patent office.
- *Priority Application Information and Date*: The original application number becomes the priority application number. Also, the original application date becomes the priority application date and it is the application date of the whole patent family.
- *Designated States*: Where applicants have requested their invention to be protected by means of a European or Patent Cooperation Treaty (PCT)[18]. Applicants have to designate the states or countries in which they want the patent to take effect and pay the appropriate fees.

As an example, in *Appendix B*, we present a complete patent family of the firm *NOVOZYMES* as offered by *DII*.

[18] PCT is an international patent law treaty that assists applicants in seeking patent protection internationally for their inventions

3.2.3 Data Collection

In general, each patent application and, in particular, each pending application is confidential until a certain stage in the proceedings (e.g. upon patent grant), or until a certain date (e.g. 18 months after filing). After that date (or that stage), the patent application is made public and anyone interested can have access to the information that is contained within the application. This procedure is followed by the most important patent offices worldwide, such as the European Patent Office (Harhoff and Wagner 2006), the World Intellectual Property Organization (The Thomson Corporation 2007), the Japan Patent Office (Kondo 1999) and the United States Patent and Trademark Office[19].

With this in mind, in December 2011, when we started to gather the data for this dissertation, we understood that the last year for which we could collect patent data was the year 2009. We could not include the year 2010 because the patents that had been applied after July 2010 would not have been made public in December 2011. Thus, the Derwent Innovation Index Database would not contain information about these patents. Even if our first data source, the *Scoreboard*, could provide us with economic data about the year 2010, our second source *DII* could not, and consequently, the upper limit of our research period was the year 2009. Concerning the lower limit, this was the economic year 2003, because the first *Scoreboard* was published in 2004 and concerned the financial year 2003.

Therefore, in December 2011, in order to build our patent-based variables, we downloaded the full information of the patent families that have been applied by the firms of our sample for the period from 2003 to 2009, by using the assignee codes that

[19] According to the website of United States Patent and Trademark Office: www.uspto.gov.

are provided by *DII*[20]. These were 142.720 patent families. Moreover, we downloaded the patent families that have cited the focal patent families until December 2011, which were 144.414, and the patent families that are cited by the focal patent families, which were 684.677 patent families.

Then, all these data were imported into a relational database management system (we chose MySQL, which is an open source database). We used a relational database management system in order to be able to retrieve easily the desirable information by using SQL queries. Moreover, we used the PHP scripting language in order to build our patent-based variable in the cases where the SQL queries were not adequate to do it by themselves. In particular, each patent-based variable has been computed by using a specific SQL query in our database and by using a special PHP script were needed.

We build five tables to store all the information of the patents provided by *DII*. The first table, labeled *Main_Patents,* contains the full information of the patent families that have been applied by the firms of our sample for the period from 2003 to 2009. The second table, labeled *Forward_Patents,* contains the full information of the patent families that have cited the patent families of the table *Main_Patents* until December 2011. The third table, labeled *Main_Forward,* connects each patent family of the table *Main_Patents* with the patent families of the table *Forward_Patents* (i.e. it connects each patent family of our sample's firms with the patent families that have cited the focal patent family). The fourth table, labeled *Backward_Patents,* contains the full information of the patent families that have been cited by the patent families of the table *Main_Patents*. Finally, the fifth table, labeled *Main_Backward,* connects each patent family of the table *Main_Patents* with the patent families of the table

[20] For example, the assignee code of the pharmaceutical firm *Sanofi-Aventis* is SNFI-C, the assignee code of the biotechnology firm *Amgen* is *AMGE-C* and the assignee code of the chemical firm *Dow Chemical* is *DOWC-C.*

Backward_Patents (i.e. it connects each patent family with the cited patent families). In *Appendix C*, the tables of our database and their characteristics are presented.

The three following figures depict the firms in our sample for each of the three industries ordered by the number of citations that their patent families, produced from 2003 till 2009, received until the December 2011. In other words, these figures present an ordering of our sample firms on the basis of the technological impact of their inventions that were developed from 2003 till 2009, as reflected in the citations these patents received in the time window between their "going public" and 2011.

Figure 3.1 Biotechnology Firms Ordered by the Number of Patent Citations Received

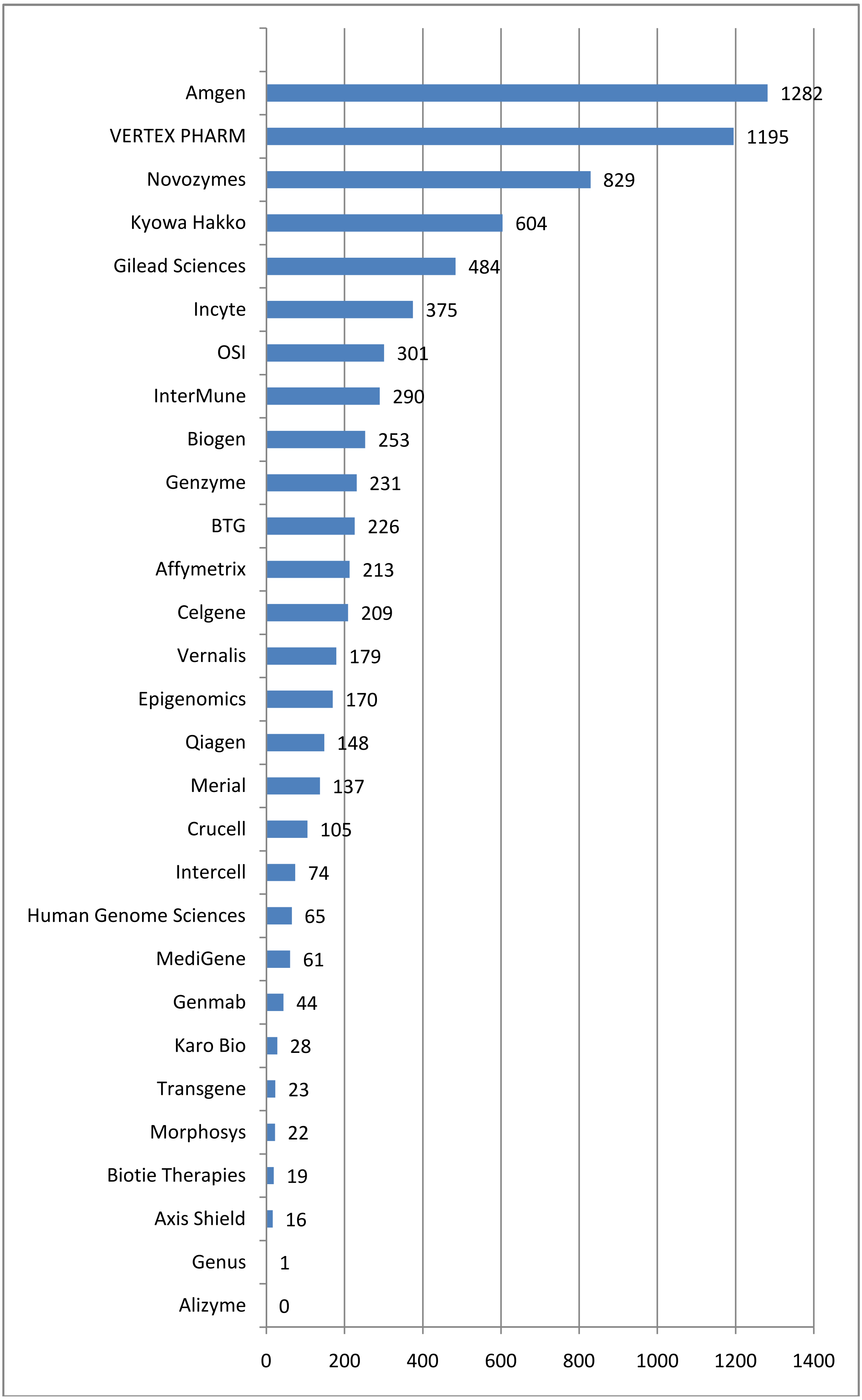


Figure 3.2 Pharmaceutical Firms Ordered by the Number of Patent Citations Received

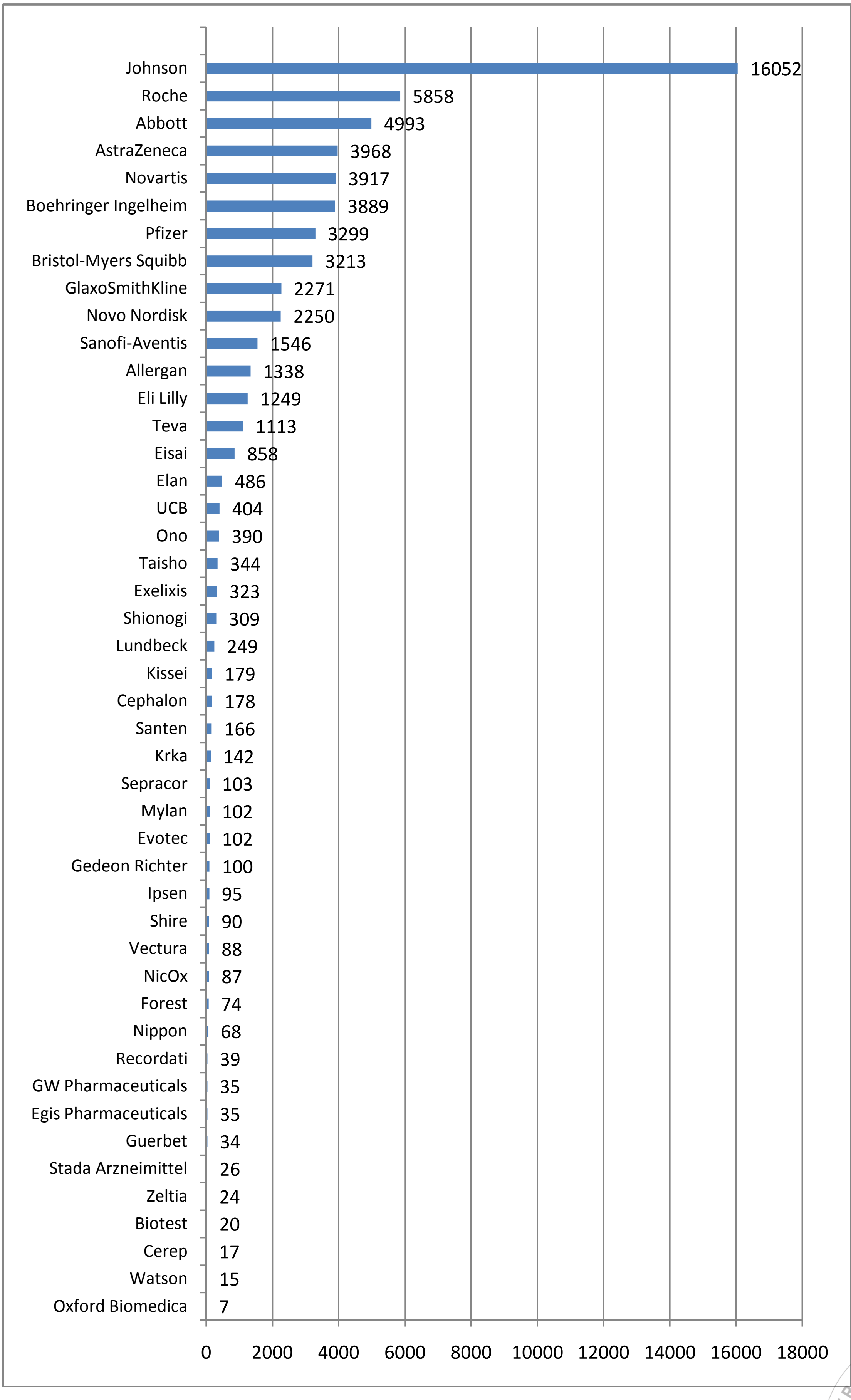
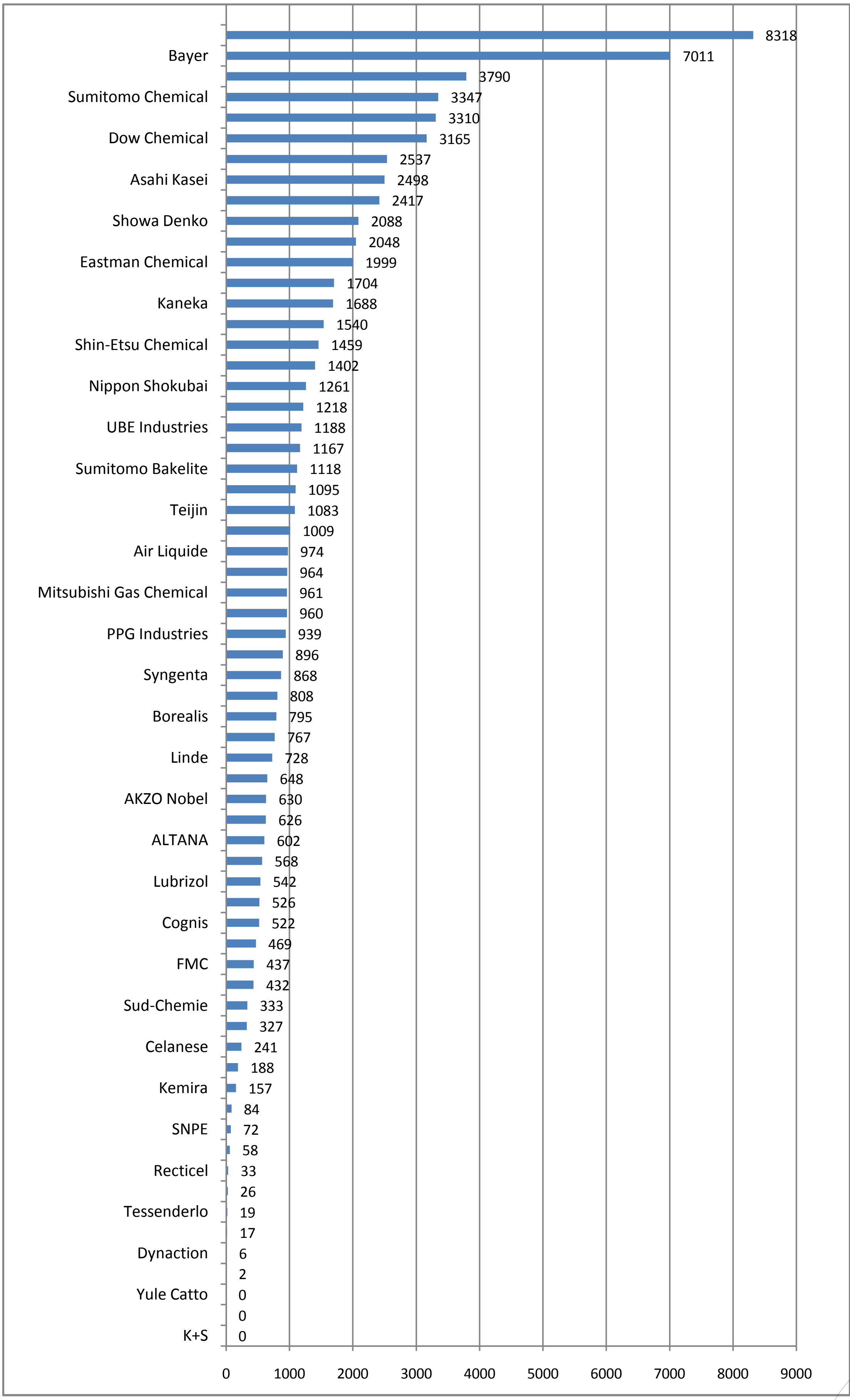


Figure 3.3 Chemical Firms Ordered by the Number of Patent Citations Received



CHAPTER 3

3.3 Variables

3.3.1 Dependent Variables

3.3.1.1 Measurement of the Impact on Technological Evolution

As we already stressed, in this study we are focusing on the factors that affect the ability of a firm to influence its environment and, in particular its technological environment. In general, when measuring the effects of a firm on its environment, it is necessary to identify and measure a specific type of *firm level actions* whose *consequences* on the environment can be also identified and measured. For example, in the case of technological environment, the patented inventions of a firm can be considered as the *firm level action* and the number of citations that these patents receive by subsequent patented inventions can be considered as the *magnitude of the inventions' impact* on technological environment[21].

More specifically, the set of all technological inventions that have been applied to patent offices worldwide can be used as a proxy for the evolution of technology. That is to say, the patented inventions can capture to a large extent (but not absolutely – see subsection *5.5* for the limitations of our study) how technologies

---

[21] Generally speaking, although scholars interested in the relationship between organizations and technological evolution have found it convenient to measure technological impact based on patent citations, this convenience does not exist in measuring the impact of organizations on different aspects of their environment, such as the institutional environment. There exist very few studies that have tried to measure the effect of organizations on evolution of institutions. For example, Hoffman (1999), trying to answer the question of who was relevant in defining legitimate environmental action for U.S. chemical industry, analyzed statistically the federal legal cases, based on the logic that whoever is participating in the legal process has a voice in determining institutional norms. Furthermore, in the same study, content analysis of a chemical journal was used to detect shifts in conceptions of the environmental issue and to discover the firms that were responsible for these shifts. In addition, Child, Lu, and Tsai (2007) applied also content analysis to various Chinese official publications concerning environmental issues, in order to identify the institutional entrepreneurs' activities that shaped the trajectory of institutional development. The lack of this kind of studies and the difficulty in constructing the metrics for computing the organizations' impact (e.g. analyzing federal legal cases, content analysis), depicts that, generally, measuring the organizations' influence on the evolution of their environment is not an easy task.

evolves step-by-step. We say step-by-step because, by using patents as a proxy of technological change, actually we split technological evolution into steps where each step is a new patent. So, the set off all steps (e.g. the set off all patents) constitute the evolution of technology.

In answer to the question "which patents have the greater impact on technological evolution", we can examine the prior art upon which each new patent is build upon. Consequently, by analyzing the citations that each new patent makes on previous patents, we can capture the most influential patented inventions. Or, in other words, by counting the number of citations that each patent receives from subsequent patents, we can compute the impact of each patent on technological evolution. And, if we want to answer to the question "which firms exerts the greater influence on technological evolution", we simply use the aggregate number of the citations of all the patents that belong to a certain firm. Therefore, by the term "firm technological impact", we refer to the impact of the firm's patents on future patents, as reflected by the aggregate number of citations received by its patents.

And, although, patents and patent citations are not the only manifestations of new technological inventions and technological impact respectively, they are probably the only manifestations that are characterized by high reliability, objectivity, and accessibility and by high time and spatial coverage; characteristics which are very important for a researcher who is searching for the most appropriate measures of technological evolution and technological impact. For example, a firm's inventor can develop inventions that are never patented and, therefore, never recorded, or he/she can be influenced by technical knowledge from sources that cannot be easily captured (e.g. from a discussion with other inventors, from not patented inventions, from products by reverse engineering, etc).

Going back to our case, the dependent variable of our primary model is the organizations' impact on technological evolution and we rely on patent citations to measure this impact. Thus, we are using the patents of a firm and the citations they receives in order to construct a firm level characteristic, that is to say, the firm's technological impact. We build this firm level characteristic by measuring the aggregate number of citations that the patents filed by a given firm in a given year have received until a given point in time (in our case this point is the December 2011, the month that we gathered the data for this dissertation). More precisely, using patent families instead of patents (since *DII* is a based on patent families), the firm's impact on technological evolution is computed as the aggregate number of citations that the firm's patent families filed in year t ($2003 \leq t \leq 2009$) received until December 2011. We must note that the citations made by the same firm that generated the patent family (self-citations) are included in our dependent variable.

Moreover, it is important to mention that we preferred this type of measure (absolute number of patent citations per year) instead of the percentage (patent citations per patent per year) because in the case of percentage we would have measured the average impact of each patent (per year) and not the aggregate technological impact of a firm (per year). For example, let use assume that the firm *A* in year *t* had applied for 10 patents which had received in aggregate 40 citations until December 2011 and the firm *B* in the same year had applied for 1000 patents which had received in aggregate 800 citations until December 2011. If we use the percentage, then the firm *A* will score 4 for year *t* while the firm *B* will score 0.8. This result indicates that the firm *A* has a 5 times greater impact on technological evolution than the firm *B*. On the other hand, if we use the absolute number of patent citations, then the firm *B* (800 citations) will appear to have a 20 times greater technological impact than firm A, an inference that is completely opposite to the latter one. But since, our research objective is to capture the overall technological impact of the inventive activity of a given firm in a given year, ***regardless*** of the number of patents in which this activity has resulted, the appropriate measure is the absolute number of patent citations and not the percentage. We are interested in the technological influence of the set of patents of a certain firm in a certain year, **seen as a unified group** and not seen as separated units of influence (e.g. patents). After all, we are focusing on the impact that **the firm** has on its technological environment and not on the technological impact of the patent.

The following examples will make more clear how exactly we compute our dependent variable. The firm *NOVOZYMES* from the biotechnology industry applied for 96 new patents families in 2003. These 96 patent families had received on aggregate 302 citations from other patent families (including its own subsequent (post-2003) patent families) until December 2011. Thus, the value of the dependent variable for the firm-year observation *NOVOZYMES-2003* is 302. The same firm in 2004 applied for 123 patents families. These patents families had received 251 citations until December 2011. Thus, the value of the dependent variable for the firm-year observation *NOVOZYMES-2004* is 251.

At this point, we will present some information with regard to patents and patent citations. Patent is a temporary monopoly awarded to a patent assignee for the commercial use and protection of a new invention (Trajtenberg et al. 1997). Patent data is the most popular indicator in invention studies (Schoenmakers and Duysters 2010) because it represents an externally validated measure of technological novelty (Ahuja 2000) and it offers important time and spatial coverage. Moreover, a patent document can provide valuable information to researchers such as the organization that generated the patent, the invention date, which is the closest link between the timing of the invention and its recording (Rothaermel and Hess 2007), the location of the invention and the technological classes in which a patent is classified (Phene et al. 2006). In addition, since each patent is been publicly revealed through the publication of the patent document after a specific period of time, patents can be viewed as codified knowledge that can be used by firms other than the originators (Hoetker and Agarwal 2007). A patent document provides information about this knowledge transfer by including the citations that a patent makes to other patents (backward citations) and the citations that the patent receives by other patents (forward citations).

All patent applications must include a list of references to all "prior art" of which the applicants are aware (Miller et al. 2007). Patent citations serve the important legal function of delimiting the scope of the property right granted to the patent (Hoetker and Agarwal 2007) and emphasizing the novel aspects of the invention (Sørensen and Stuart 2000). The integrity of the citation procedure is maintained by patent examiners, who guarantee that relevant patents will be cited and irrelevant patents will be omitted (Stuart 1998, Hoetker and Agarwal 2007). As we showed earlier in *Table 2.1*, researchers in management science have deployed patent citation data in multiple ways conditional on their view and on their scientific objectives (e.g. the nexus between breakthrough inventions and large established corporations, the linkage of basic science and technological impact, the structure of organizational knowledge bases, technological collaboration networks, R&D consortia and more).

Rothaermel and Boeker (2008) noted that "patent citations can be viewed as 'technological fossils' representing the intellectual lineage of new patents" while Stuart (2000) stressed that "just as citations between journal articles reveal the transmission of ideas between papers, patent citations trace technological ancestries". Moreover, Rosenkopf and Nerkar (2001) stated that "when other firms recognize and build upon a firm's knowledge, it demonstrates this firm's influence on the overall evolution of a particular product class or technology" and measured that influence by using patent citations.

A reasonable comment on the measure of the impact on technological evolution could be that we must take into account only the patent inventions that have been granted by a patent office and not all the patents that have been applied for. However, we must have in mind that after a specific period of time from the date of application (usually after 18 months), the patent application is disclosed, even if the decision for the patent is still pending. Therefore, this patented invention can have an impact on technological evolution since the novel knowledge that is claiming for has been revealed and the future patents can rely on and cite this patent.

Finally, we have to mention that, in order to avoid repetition, in the remainder of the document, phrases such as "magnitude of impact", "technological impact", "technological effect" or "technological influence" refer literally to the impact on the technological evolution.

3.3.1.2 Measurement of the Breadth of Technological Impact

In this study, in order to capture the different aspects of the breadth of a firm's technological impact, we use three distinct measures of breadth.

The first measure regards the technological impact of a firm on industries other than its industry. Rosenkopf and Nerkar (2001) examining the impact of firms in the optical disk industry, named this type of breadth "overall impact" and computed it as the total number of citations from non-optical disk patents received by the patents of the firms of the optical disk industry. The same variable, but with different name ("nondomain impact") was computed also by Miller et al. (2007).

In our case, we name this type of breadth as *Beyond-Industry Impact* and we compute it as the total number of citations that a firm's patent families filed in year t (2003 ≤ t ≤ 2009) received until December 2011 from patent families that are classified in international patent classes (i.e. the technological classes in the patent system) that are different from those of the focal firm's industry. The international patent classes (IPC) of a particular industry (i.e. pharmaceuticals, biotechnology or chemicals) are defined as the set of all the distinct international patent classes of the last three year's patent families of the industry's firms[22]. Needless to say, that the set of international patent classes of each industry changes every year since it is actually a three-year moving window. So, we construct the technological classes of each industry based on the international patent classes of each patent family of each firm of this industry (until three years back). Thus, for example, even if a particular patent family of a pharmaceutical firm is categorized in a typical chemical international patent class, this typical chemical international patent class will be included in the set of the pharmaceutical industry and not in the set of the chemical industry, because the firm belongs to the pharmaceutical industry. Consequently, each industry is defined on the basis of the prior patented technologies of its members and not on the basis of the technologies *per se*. By using this measure, we achieve to measure the impact on technologies that are not related to the technologies of a particular industry.

As an example, the firm *NOVOZYMES* applied for 96 new patents families in 2003. These 96 patent families had received in aggregate 302 citations from other patent families until December 2011. From these 302 patent families, 16 patent families are categorized in international patent classes that are not included in the set

[22] We chose three years because by choosing less than three years we run the risk of excluding technological classes in which a firm is probably still active although it hasn't completed the development of any patent in those classes. On the other hand, by choosing more than three years we run the risk of including technological classes in which a firm is inactive.

of all the international patent classes of all the last three years' patent families (from 2001 to 2003) of all sample firms in the biotechnology industry. As was the case with the measurement of technological impact, we chose the absolute number of patent citations and not a percentage because we are interested in the firm's breadth of impact and not in the patent's breadth of impact.

At this point, we have to mention that the international patent classes (IPC) are the technical areas to which the patent examiner assigns each patent, after it has been applied (Nemet and Johnson 2012). Since we are working with patent families and not with simple patents and because each patent family may consist of multiple patents and each patent may belong to multiple international patent classes[23], we defined the most frequent four-digit IPC as the representative IPC of the patent family. So, when we refer to the international patent class of a patent family, we refer to the most frequent four-digit international patent class of the patent family. To explain more, the first four digits of the IPC indicate the generalized technological class while the rest digits indicate the specialized technological class. We assume that two patent families with the same four-digit IPC largely belong to the same technological area.

The use of the most frequent four-digit international patent class of a patent family is a very convenient way (computationally) when trying to answer the question "what is the technological class of a certain invention in terms of IPC?". Instead of answering with a list of all the international patent classes and their presences or not in each patent of a patent family, we simply attribute a certain four digit IPC to each patent family as a representative IPC. For example, a certain patent family of the biotechnology firm NOVOZYMES coded as 2005-758242 from the *DII* consists of 20

[23] As we mentioned earlier, patents in the "patent family" are either the same invention applied for protection in different markets or slight improvements of the same invention. So, in the case of the slight improvements, it is possible that patents of the same patent family to include different IPC.

different patents. All the distinct international patent classes that are included in these patents are 30 and follow: C12N-000/00, G01N-015/00, C12Q-001/68, G01N-033/48, C07H-021/00, C07H-021/04, C12N-001/08, C12N-009/00, C12N-015/00, C12N-015/11, C12N-015/87, C12P-021/04, C40B-030/04, C40B-040/04, C40B-040/08, C12N-015/57, C12N-015/74, C12N-009/42, C12N-009/24, C12N-015/56, C12N-015/75, C07K-014/32, C07K-014/195, C12N-009/10, C12N-009/12, C12N-009/16, C12N-009/56, C12N-009/78, C12N-009/88, C12N-009/90. The frequencies of all four-digit IPC are: C12N – 19 times, C40B – 3 times, G01N – 2 times, C07H – 2 times, C07K – 2 times, C12Q – 1 time, C12P 1 time. Therefore, the four-digit IPC C12N, which is the most frequent IPC, can represent the technological class of the patent family. And so, with regard to the set of the technological classes of a certain industry, this set consists of all the representative international patent classes of each patent family of each firm of the industry until three years back.

The second measure of the breadth of technological impact is the *Beyond-Firm Impact*. With this type of impact, we attempt to capture the degree to which a firm exerts technological influence on firms that operate in different technological areas than those of the focal firm. This measure is like *Beyond-Industry Impact* but instead of focusing on the impact on patent families that are not related to the technological fields of a particular industry, we focus on the impact on patent families that are not related to the technological fields of a particular firm, regardless of their industry. That is to say, *Beyond-Firm Impact* attempts to capture the effect of a given firm on other firms with which they do not share any common specialized area of research, even if these two firms are listed in the same industry. For example, when two firms from the biotechnology industry operate in completely different specialized technological fields and the inventions of one firm affect the inventions of the other, this impact is regarded as *Beyond-Firm Impact.*

Apparently, *Beyond-Firm Impact* is a less wide (or of the same wide at most) measure of impact compared to *Beyond-Industry Impact*. Or, in other words the *Beyond-Industry Impact* endeavors to capture the extreme cases of breadth of impact where a patent family influences other patent families that are located exclusively in very distant technological areas, while *Beyond-Firm Impact* endeavors to capture in general the breadth of technological impact, by measuring the influence on patent families that are categorized in technological fields that are different from those that the focal firm operates in (regardless of the technological fields' industries). By using these two different measures of breadth impact, we attempt to distinguish those firms that exert broad influence on technological evolution in general, from those firms whose breadth of technological impact is so wide that their inventions are found to be useful even in very distant, beyond-industry technological fields.

We compute *Beyond-Firm Impact* as the absolute number of citations that a firm's patent families filed in year t ($2003 \leq t \leq 2009$) received until December 2011 from patent families that are classified in international patent class which are different from the set of all international patent classes of the focal firm. The set of all international patent classes of a focal firm is defined as the total of all the distinct international patent classes of the last three year's patent families of the focal firm. By using this measure, we achieve to measure the impact of a firm on technologies that are not related to the technologies that have been developed within the firm's R&D department.

As an example, the firm *NOVOZYMES*, as we mentioned previously, applied for 96 new patents families in 2003 which received in aggregate 302 citations from other patent families until December 2011. From these 302 patent families, 35 patent families are categorized in international patent classes that are not included in the set of all the international patent classes of all last three years' patent families (from 2001 to 2003) of the firm *NOVOZYMES*. Or, in other words, none of the international patent classes of these 35 patent families are included in the set of the international patent classes of the focal firm's patent families applied from 2001 to 2003.

We remind that when we refer to the international patent class of a particular patent family, we refer to the most frequent four-digit international patent class of the patent family.

Finally, the third measure of breadth of impact regards the ***concentration*** of the international patent classes of the patent families that are affected by a firm's patent families. This measure endeavors to capture the extent to which a firm develops inventions whose influences are spread across different technological fields (Trajtenberg et al. 1997). We name this measure *Herfindahl Type Breadth* since it is based on the Herfindahl Index. *Herfindahl Type Breadth* differs from the first two measures (*Beyond-Industry Impact, Beyond-Firm Impact)* in that it measures the distribution or the concentration of the international patent classes of the affected patent families and not the absolute numbers of the affected patent families as the first two measures. By employing this measure, we want to examine if the influenced patent families are concentrated into a narrow or into a wide range of technological classes.

Valentini (2012) used this type of breadth of impact as a dependent variable to examine the effect of mergers and acquisitions on patenting quality, while Argyres and Silverman (2004) used it to explore the link between a firm's R&D structure and the breadth of its innovations' impact.

In order to construct this variable, we first measure the *Herfindahl Type Breadth* at the patent level of analysis. We define *Herfindahl Type Breadth* for **patent family *i*** as:

$$1 - \sum_{k=1}^{k} \left(\frac{NCITING_{ik}}{NCITING_i}\right)^2 ,$$

where *NCITING* represents the number of patents families that cite the focal patent family and k indexes the international patent classes of the citing patent families. As mentioned earlier, because each patent family may consist of multiple patents and each patent may belong to multiple international patent classes, we defined the most frequent international patent class as the representative international patent classes of the patent family. The intuition behind the third measure of the breadth of technological impact is that the more diverse the technological classes of the citing patents, the broader and the more general the impact of the cited patent (Argyres and Silverman 2004). Therefore, a high *Herfindahl Type Breadth* score indicates that an invention has affected a wide range of technologies (Hoetker and Agarwal 2007). To measure this variable on a **firm-year base**, we computed the average score of *Herfindahl Type Breadth* of the patent families of a given firm in a given year. As an example, in year 2003, each patent family of the firm *NOVOZYMES* had on average a *Herfindahl Type Breadth* score of 0.8[24].

Finally, because, in general, Herfindahl type measures suffer from bias due to the properties of small numbers (e.g. a patent that receives either a single citation or two citations but from the same international patent class will score zero in

[24] For comparison, we note that if, for example, a patent family receives citations from 10 patent families that are categorized in 10 different IPC, the value of the Herfindahl Type Breadth of this patent family would be: 1 – ($1^2/10^2$ + $1^2/10^2$ + $1^2/10^2$ +$1^2/10^2$ +$1^2/10^2$+$1^2/10^2$+$1^2/10^2$+$1^2/10^2$+$1^2/10^2$+$1^2/10^2$)= 1- (10/100)=90/100=0,90. But if all the 10 patent families are categorized in the same IPC, then the value of the Herfindahl Type Breadth would be: 1 – ($10^2/10^2$)=0.

Herfindahl Type Breadth), we consider only the patent families that receive over three citations, in an effort to lessen this particular bias without losing a significant amount of observations (Argyres and Silverman 2004). To explain in more detail this type of bias, we give the following example: Firm *A* applied for three patent families in year *t* and each patent family has received only one citation. Because there exists only one citation, the *Herfindahl Type Breadth* for each patent family is zero and so, the average *Herfindahl Type Breadth* of firm *A* in year *t* is zero. In addition, firm *B* applied also for three patent families in year *t* and each patent family has received 100 citations from patent families that belong to the same IPC. For this case, the score of each patent family is zero and the average *Herfindahl Type Breadth* of firm *B* in year *t* is zero as well. However, it is unfair to consider firm *B*, which developed inventions characterized by a "proven" narrow breadth of impact, of equal value with firm *A*, for which it is quite irrational to come to conclusion regarding the breadth of its inventions' impact since they have received only one citation each. Therefore, the more the number of patent citations approaches zero, the less the value of *Herfindahl Type Breadth* reflects the reality. Thus, in order to avoid such phenomena, we take into account only the patent families that receive over three citations.

To recap, we use three distinct measures of the breadth of technological impact. The first (*Beyond-Industry Impact*) focuses on the firm's impact on industries other than its own, the second (*Beyond-Firm Impact*) focuses on the impact on technological areas different from the areas that the focal firm operates in, and the third (*Herfindahl Type Breadth*) focuses on the ***concentration*** of the international patent classes of the affected patent families. The first two measures are integers and are based on the absolute values of the number of the influenced patent families while the third measure is a value between 0 and 1 and measures distribution. All three measures can be considered as different and interesting aspects of the breadth of a firm's impact on technological evolution and, hence, all the three dimension worth examining.

3.3.2 Independent Variables

Scientific Basis. We use two proxies for capturing the extent to which firms build on scientific knowledge. Following a well-established practice (Operti and Carnabuci 2014), we use *Non Patent References* as the first proxy for *Scientific Basis*. *Non Patent References* refer to all patents' references that are not patent documents. Those references can be scientific papers, technical papers, conference proceedings, textbooks, disclosure bulletins, and so forth (Meyer 2000). To compute this measure, for each member of the patent family, all the distinct (to guard against double counting) *non patent references* have been collected, in order to establish the number of *non patent references* that each patent family cites (Sapsalis et al. 2006). Then for each firm-year observation, we computed the average number of the patent family's *Non Patent References*. As an example, each patent family of the firm *NOVOZYMES* in year 2003 cited on average 13.6 *Non Patent References* while in 2004 the patent families of the same firm cited on average 9.7 *Non Patent References*.

The second proxy for the *Scientific Basis* is the *Scientific Inventions*. *Scientific Inventions* refer to those patented inventions that have been applied in collaboration with at least one university. Firms that collaborate to a large extent with universities for developing new inventions may be considered as firms whose inventions rely heavily on scientific knowledge. To compute *Scientific Inventions*, for each firm-year observation, we used the ratio of the patent families that have been applied by the firm in cooperation with a university to total patent families. As an example, in year 2003, the firm *NOVOZYMES* applied for 2 patent families in collaboration with a university. And, since its overall patent families were for that year 96, the ratio of *Scientific Inventions* was 0.021.

Technological Collaboration. We employed the number of patent co-assignees to measure *Technological Collaboration*. Each organization[25] which is involved in the development of an invention will demand the same rights over the patent with the rest of the partners. Those rights can be secured by including the involved organization in the patent assignee list. Thus, for each firm-year observation, we computed the average number of co-assignees of each patent family. As an example, in year 2003, the firm *NOVOZYMES* applied for 96 patent families and each patent family had 1.71 co-assignees on average.

Technological Diversity. Following Nerkar and Shane (2007) and Leone and Reichstein (2012), we measured *Technological Diversity* using the number of international patent classes (IPC) that each patent document has on its front page. We remind readers that IPC are the technical areas to which the patent examiner assigns each patent, after it has been applied (Nemet and Johnson 2012). IPC have been commonly used as indicators of invention's technological breadth and diversification (Khoury and Pleggenkuhle-Miles 2011). We name this measure *IPC Diversity*. To construct *IPC Diversity*, for each firm-year observation, we computed the average number of the distinct four-digit IPC of each patent family. As an example, in year 2003, each patent family of the firm *NOVOZYMES* was assigned to 3.74 four digit IPC on average.

[25] A patent assignee is usually an organization (firm, university etc) but it can also be an individual. We make no distinction between organizations and individuals when measuring the number of patent co-assignees.

In addition, we used the *Originality* indicator, proposed by Trajtenberg et al. (1997), as a second proxy for *Technological Diversity* (Banerjee and Cole 2010). *Originality*, which is based on the Herfindahl Index, is a citation-based index which measures the dispersion of a patent's ***backward*** citations (i.e. citations made by a patent) across different technological domains (Lettl et al. 2009). *Originality* is defined for **patent family *i*** as:

$$1-\sum_{k=1}^{k}\left(\frac{NCITED_{ik}}{NCITED_{i}}\right)^{2},$$

where *NCITED* represents the number of patents families cited by a focal patent family and k indexes the IPC of the cited patent families. As was the case with the *Herfindahl Type Breadth*, we defined the most frequent four-digit IPC as the representative IPC of the patent family. A high *Originality* score indicates that an invention has drawn upon a wider range of technologies (Hoetker and Agarwal 2007). To measure this variable on a firm-year base, we computed the average score of *Originality* of the patent families of a given firm in a given year. As an example, in year 2003, each patent family of the firm *NOVOZYMES* had in average an originality score of 0.79.

New-to-Firm Technological Domains. In order to measure *New-to-Firm Technological Domains*, we based again on the international patent classes of each patent family. Influenced by the work of Ahuja and Morris Lampert (2001), we consider that a firm is entering a new technology domain when it applies for a patent in an IPC in which it had not applied in the previous 4 years. This consideration is based on the assumption that technological classes in which a firm has not been active in the last 4 years are considered to be novel for it. Thus, for constructing this variable, for each firm-year observation, we measured the number of patent families that include at least one IPC that is new to the firm. As an example, in year 2003, the firm *NOVOZYMES* applied for eight patented inventions that were categorized in new-to-firm international patent classes. Therefore, the value of the variable *New-to-Firm Technological Domains*, for firm *NOVOZYMES* in year 2003 is 8.

Unexplored Technological Areas. We follow previous studies in measuring *Unexplored Technological Areas* by measuring the "shortage" of citations to prior patent families made by a focal patent family (Sapsalis et al. 2006, Operti and Carnabuci 2014). In particular, Operti and Carnabuci (2014) by counting the average number of backward citations a firm's patents make to previous patents measured the extent to which a firm "cumulatively build on a body of existing knowledge or, conversely, experiment with pioneering technologies". In our case, for each firm-year observation, we computed the average number of citations to prior patent families each patent family makes[26]. The logic behind this measurement is the following: a lack of references to prior patents reveals the development of an invention that has only a few technological antecedents, and thus, it indicates the inventive activities of the firm in technological areas that are largely unexplored. As an example, in year 2003, each patent family of the firm *NOVOZYMES* cited on average 10.78 patent families.

However, it is important to mention that Ahuja and Lampert (2001) refer to the experimentation with unexplored technological areas as "Pioneering Technologies" and they measure it in a different way. They measure it as a dummy variable (1 = no backward citations; 0 otherwise), while we measure it as the average number of citations a firm's patents make to previous patents , following other studies such as Sapsalis et al. (2006) and Operti and Carnabuci (2014). The main advantage

[26] To avoid double counting, when two patents from the same patent family cite the same patent or cite two different patents that belong to the same family, we count it as one citation.

of the measure that we employ is that it distinguishes between patents that cite just a few prior patents from patents that cite a lot of prior patents and, so, it accounts for the quantity of the backward citations. The main advantage of the measure of Ahuja and Lampert (2001) is that it can capture the pure “Pioneering Technologies” that have zero technological antecedents. The major reason for our choice is that it is very rare in our sample to find a patent family with zero backward citations and so, if we had follow Ahuja and Lampert (2001), the value of *Unexplored Technological Areas* would have been almost fixed (e.g. zero) and the regression models would have been unable to trace the effect of this particular variable on our dependent variables (since almost all the observations would have the same value).

Finally, we have to mention that we reversed the sign of this variable before being entered into our econometric models in order the positive estimated coefficients to agree with the confirmation of the hypothesis and not with the rejection. Therefore, after that reversal, an increase in this explanatory variable shows a reduction of the backward citations’ number, which, in turn, denotes an increase in the experimentation with unexplored technological areas. Such reversal has no other effect on the regression models except for the sign of the estimated coefficients of the focal variable.

Recency of Technological Inputs. Following previous research (Rosenkopf and Nerkar 2001, Cattani 2005), we employed the application year of the cited patent families to compute this variable. More specifically, for each patent family, we computed the average time difference between the application year of the focal patent family and the application year of all the distinct cited patent families. We did the same for all patent families each firm applied for in a given year. We assume that this difference in age can adequately capture the degree to which a firm experiments with leading-edge technologies. As an example, in year 2003, each patent family of the firm *NOVOZYMES* cited previous patents that were, on average, 9.9 years old.

Moreover, it is important to note that, as was the case with the *Unexplored Technological Areas*, we reversed the sign of this variable before we included it into our regression models in order the positive estimated coefficients to confirm the hypothesis and not to reject it. We did this because *Hypothesis 6a* and *6b* is build on the recency of technological inputs but our metric actually measures maturity (the average time difference between the application year of the focal citing patent family and the application year of all the distinct cited patent families).

Recent Technological Activity. In the realm of firms with high R&D expenditures, the number of patents of the previous years can adequately indicate how active a firm was in terms of new technological developments. Most of the studies have used a time period of five years to measure *Recent Technological Activity* (Phene et al. 2006, Vanhaverbeke et al. 2009, Srivastava and Gnyawali 2011), while some authors have used a time window of three years (Nerkar 2003, Hoetker and Agarwal 2007). Furthermore, the extant literature has labeled this variable with various names such as patent stock (Phene et al. 2006), technological capital (Vanhaverbeke et al. 2009), prior capabilities (Nerkar 2003), and knowledge base size (Operti and Carnabuci 2014) while the term *Recent Technological Activity* has been used by Hoetker and Agarwal (2007). We computed *Recent Technological Activity* as the logarithm – according to Wooldridge (2002) taking logs is a standard rule of thumb for large integer values – of the sum of the patent families the firm had applied for in the prior five years, for each firm-year observation. As an example, the *Recent Technological Activity* of the firm *NOVOZYMES* in year 2003 was 424 (2.63, in logs), which means that *NOVOZYMES* in the period from 1998 to 2002 had applied for 424 patent families.

Internal Focus. As prior research suggests, the extent to which a firm builds upon its own technological developments is manifested by the frequency with which it cites its own patents inventions (De Carolis 2003, Hoetker and Agarwal 2007). In the present study, for each firm-year observation, *Internal Focus* is measured as the average ratio of the citations that each patent family makes to prior patent families held by the focal firm to its total backward citations. As an example, in year 2003 on average, 26% of the backward citations of each patent family of *NOVOZYMES* were citations made to its own prior patent families.

R&D Investment. We used the log of yearly R&D expenditure to measure *R&D Investment* (Cattani 2005). Since, R&D investment creates patent applications always with a time-lag (Kondo 1999) we used the one-year lagged value. Moreover, we included also the contemporaneous value because it is possible that the new R&D projects which start at the beginning of the year to result in patent application at the end of the same year. Finally, even if in the cases of high-impact inventions one would expect that R&D activities would need a lot of time to materialize into something concrete and novel, we decided not to include more than 1-year lags because we would lost a lot (139 – as many as the firms of our sample - observations per year) and "important" observations. To explain clearly what we mean by the adjective "important", we remind readers that the number of patent citations that the patent families of the year 2003 have received until December of 2011 indicates much better the extent of the patents' impact as compared to the patent families of the year 2009. There is a larger period of time for early inventions (e.g. 2003) to "show" their influence in comparison with the late inventions (e.g. 2009) for which there is not enough time to "prove" their degree of influence. This is the reason why we characterize the early inventions of our sample as more important than the late inventions.

Firm Size. We operationalized *Firm Size* as the natural logarithm of the number of firm employees (Yayavaram and Ahuja 2008).

3.3.3 Control Variables

We consider as control variables all the patent characteristics that affect the number of forward citations and, at the same time, they cannot be deployed reliably as proxies for specific firm characteristics. For example, *Patent Co-assignees* is a patent characteristic that can be used as a proxy for the extent to which a firm collaborates in order to develop new technological inventions. On the contrary, it is very difficult to attribute a certain firm characteristic based on the *Number of Inventors* of each patent family. Apart from the patent control variables, we include also as a control variable the firm profitability. Although one would expect that this variable should be examined in the hypotheses section, we chose to consider it as a control variable. The reason is that such variable as profitability is too generic since it encompass too many, too diverse factors to be seriously used to test hypotheses[27]. That is the reason we treat it as control variable.

Number of Patents. In this study, we endeavor to shed some light on the firm characteristics that enable a firm to become a significant contributor to technological evolution. In order to achieve unbiased estimations of the explanatory power of these

[27] The following plausible chain of effects, *namely higher profitability->higher R&D expenditure->more patents->more patent citations* although it is logical to be expected, it cannot be revealed in our models since we include R&D expenditure as an explanatory variable.

firm characteristics, we need to control for the number of patents in both models (i.e. technological impact and breadth of impact). In other words, we want to find the factors that make one firm more influential than another, when both of the firms have applied for the same number of patents. The reason is that the number of citations the patents of a firm receive is expected to be strictly related to the number of patents the firm has applied for (Rosenkopf and Nerkar 2001). To compute this variable, we measured the logarithm of the average number of patent families, for each firm-year observation. As an example, in year 2003, the firm *NOVOZYMES* applied for 96 patent families.

Number of Claims. Each patent contains a set of claims which are the list of the specific technological developments for which the patent assignee is claiming exclusive rights (Harhoff and Wagner 2006). Patent claims actually declare the specific novelties that are claimed to have been achieved by a particular patent family (Markman et al. 2004). With regard to the number of claims, this may reflect the breadth of the invention or just how detailed the protected area is defined (Van Zeebroeck et al. 2009).

As far as the technological impact is concerned, prior research has shown that the number of claims is positively related to the number of citations that a patent receives (Nerkar 2003, Nemet and Johnson 2012). With respect to the breadth of impact, we think that the number of claims can affect it in a positive manner. In particular, as the number of claims increases, the number of the distinct (and obviously different) technological achievements increases, and, thus, the breadth of the patent family increases as well (Van Zeebroeck et al. 2009). And as the breadth of the distinct technological developments that are included within a patent family expands, it is logical to expect also the breadth of the patent family's impact to expand.

Accordingly, to account for these potential effects, we included the variable *Number of Claims* as a control variable. To compute this variable, for each firm-year observation, we measured the average number of claims of each patent family. As an example, in year 2003, each patent family of *NOVOZYMES* included in average 5.09 claims.

Number of Inventors. Number of inventors is the number of the individual inventors[28] that are listed on a patent application and reflects the size of the team involved in the research project the output of which is being patented (Hoetker and Agarwal 2007). With regard to the technological impact, a considerable body of literature has acknowledged the number of inventors as a positive factor for the number of citations that a patent receives (Gittelman and Kogut 2003, Sapsalis et al. 2006). As far as the breadth of impact is concerned, when more inventors are involved in a patented invention, the likelihood that inventors belong in diverse technological fields increases; thus, this particular invention refers to more technological classes, on the patents of which this invention may have an impact. Consequently, we included the control variable *Number of Inventors* to account for these possible effects. This variable is computed as the average number of inventors of each patent family, for each firm-year observation. As an example, in year 2003, each patent family of the *NOVOZYMES* had been developed in average by 4.84 inventors.

Geographical Coverage. Geographical coverage refers to the geographical extension of protection of an invention's intellectual property rights. It is logical to expect that the broader the geographical coverage of an invention, the higher its

[28] Within the patent system, an inventor is always an individual and never an organization.

impact on subsequent inventions, since it can potentially affect a larger number of patents. We employed two proxies for this control variable.

The first proxy is the *Patent Family Size*. As noted previously, a patent family is a set of patents that are related to the same invention. The size of a patent family actually declares the number of the patents a firm issued in different countries for a specific invention plus the number of the patents that are considered as extensions (or improvements) of this invention (Gittelman and Kogut 2003). With regard to its relation to the technological impact, there exist several studies that have empirically demonstrated a positive effect of the size of a patent family on technological impact (Gittelman and Kogut 2003, Lettl et al. 2009). Concerning the breadth of impact, we anticipate the breadth of a patent family's impact to be affected positively by *Patent Family Size*, on the assumption that the more the patents that are included within a patent family, the more the patents that concern extensions or improvements; and, as the number of patents that concern extensions increases, the number of the distinct technological classes that are included within the patent family increases as well (the extension may concern a new technological class). And, consequently, as the technological domains of the invention increases, the technological domains that this invention can have an impact also increase. Thus, to account for the potential explanatory power of *Patent Family Size*, we obtained the average number of patents of each patent family, for each firm-year observation. As an example, in year 2003, each patent family of the firm *NOVOZYMES* included in average 6.98 patents.

The second proxy for the *Geographical Coverage* is the *Designated States*. *Designated States* concerns the number of states or countries that the patent applicants have designated as places in which the patent to take effect. We expect patent families with many designated states to be more influential, because the pool of inventions worldwide that can be potentially affected by these patent families is greater. This proxy is computed as the logarithm of the average number of designated states of each patent family, for each firm-year observation. As an example, in year 2003, each patent family of the firm *NOVOZYMES* was protected in average 100.38 states or countries.

Number of Granted Patents. Since our dependent variable is build upon the citations of the patents that have been applied for and not upon the citations of the granted patents, it is necessary to control for the number of granted patents. It is logical to expect patents that have been granted to be more influential than patents whose applications have been rejected. To compute this variable, we measured the ratio of the number of the patent families that have at least one granted patent as their member to the aggregate number of patent families, for each firm-year observation[29]. As an example, in year 2003, 45 of the 96 patent families (47%) of the firm *NOVOZYMES* included at least one granted patent. Thus, the value of this variable for NOVOZYMES in year 2003 is 0.47.

Number of Citations Received. This control variable regards only the models of the breadth of technological impact and, in particular, the case when the breadth of impact is measured as a Herfindahl index (Banerjee and Cole 2010, Valentini and Di Guardo 2012). We include this control variable in order to examine if and how *Herfindahl Type Breadth* is related to the number of forward citations (i.e. citations received).

Profitability. Generally, firms that produce profits may have the slack to invest more in R&D, and, thus, to develop more patents and to receive more patent citations. But since we will account for R&D investment, this particular effect will be

[29] A special symbol at the end of each patent number (e.g. *A1* or *A*) denotes if this patent has been granted or not.

diminished. However, profitability may still affect technological impact but in different ways, as two studies have demonstrated. In particular, Yayavaram and Ahuja (2008) controlling for the effects of firm performance (in terms of profitability) on the number of citations that a firm's patents receive, found profitability to be a significant predictor of the number of patent citations. In addition, Ahuja and Lampert (2001) included the variable *Net Income* as a control variable in a regression model in which the dependent variable was the number of patent citations and they found *Net Income* to have a significant positive effect on patent citations. Both of these studies included R&D expenditure as an explanatory variable. We used the log of the yearly *operating* profit to measure *Profitability*. As in the case of *R&D investment*, apart from the contemporaneous value, we used the one-year lagged value because the effect of profitability on technological impact happens with a time-lag.

Year. All of the models we report include year dummies. Year dummies control for all the time-varying factors that affect all firms, including macroeconomic conditions, as well as any trends in patenting rates and citation rates (Yayavaram and Ahuja 2008). Especially for the citation rates, controlling for time is very important because as a patent gets older, the number of its forward citations, *ceteris paribus*, increases (Sapsalis et al. 2006). So, each patent family is categorized in specific age cohorts (2003 age cohort, 2004 age cohort, … 2009 age cohort) and comparisons are made only for the patent families of the same cohort. We chose the application year of the patents instead of the grant year because the application year is closer in time to the time of the invention. Therefore, for the period from 2003 to 2009, we inserted annual time dummies for each year, with 2009 being the omitted year and thus serving as the reference year (Rothaermel and Hess 2007).

Country. We acknowledge that differences may exist across countries in patenting propensity and in research productivity which may have an impact on patent citations (Yayavaram and Ahuja 2008). Therefore, our models include dummy variables for the countries of origin of each firm (e.g. country dummies for USA, EU, and Japan).

Industry. Finally, since there are substantial differences across industries in terms of citation propensity, patenting propensity, technological opportunity, and invention speed (Ahuja and Morris Lampert 2001, Leone and Reichstein 2012, Nemet and Johnson 2012), we had to include dummies for the industry effect (e.g. industry dummies for biotechnology, pharmaceuticals, and chemicals) .

The following table (*Table 3.2*) illustrates all the variables of our study accompanied by their measures and their sources. ***All the variables used are time-variant*** except for the industry and the country dummies.

Table 3.2 The Set of our Variables

Variable	*Measure*	*Source*
Technological Impact	The number of citations that the firm's patent families filed in year t (2003 ≤ t ≤ 2009) received until December 2011	*DII* (List of *Citing Patents*)
Breadth of Impact – *Beyond-Industry Impact*	The number of citations that the firm's patent families filed in year t received until December 2011 from patent families that are classified in international patent classes which are different from the total of all international patent classes of the focal firm's industry	*DII* (List of *Citing Patents*)

Variable	*Measure*	*Source*
Breadth of Impact – *Beyond-Firm Impact*	The number of citations that the firm's patent families filed in year t received until December 2011 from patent families that are classified in international patent classes which are different from the total of all international patent classes of the focal firm	*DII* (List of *Citing Patents*)
Breadth of Impact – *Herfindahl Type Breadth*	The average score of *Herfindahl Type Breadth* of the patent families of a given firm in a given year. The *Herfindahl Type Breadth* for patent family *i* is defined as: $$1 - \sum_{k=1}^{k} \left(\frac{NCITING_{ik}}{NCITING_i}\right)^2 ,$$ where *NCITING* represents the number of patents families that cite the focal patent family and k indexes the international patent classes of the citing patent families.	*DII* (List of *Citing Patents*)
Scientific Basis – *Non Patent References*	The average number of the patent family's *non patent references*, for each firm-year observation (references to scientific papers, technical papers, conference proceedings, textbooks, disclosure bulletins, etc.)	*DII* (List of *Patents Cited by Inventor / Examiner*)
Scientific Basis – *Scientific Inventions*	The average number of patent families that have been applied by the firm in cooperation with a university, for each firm-year observation	*DII* (*Patent Assignee Names and Codes*)
Technological Collaboration	The average number of co-assignees of each patent family, for each firm-year observation	*DII* (*Patent Assignee Names and Codes*)
Technological Diversity – *IPC Diversity*	The average number of the distinct four-digit IPC of each patent family, for each firm-year observation	*DII* (*Patents Details* → *Main IPC*)

Variable	*Measure*	*Source*
Technological Diversity – *Originality*	The average score of *Originality* of the patent families of a given firm in a given year. The *Originality* for patent family *i* is defined as: $$1-\sum_{k=1}^{k}\left(\frac{NCITED_{ik}}{NCITED_i}\right)^2,$$ where *NCITED* represents the number of patents families cited by a focal patent family and k indexes the IPC of the cited patent families.	*DII* (List of *Patents Cited by Inventor / Examiner*)
New-to-Firm Technological Domains	The number of patent families that include at least one IPC that is new to the firm (i.e. IPC in which the firm had not applied in the previous 4 years), for each firm-year observation	*DII* (*Patents Details* → *Main IPC*)
Unexplored Technological Areas	The average number of citations to prior patent families each patent family makes, for each firm-year observation	*DII* (List of *Patents Cited by Inventor / Examiner*)
Recency of Technological Inputs	The average time difference between the application year of the focal citing patent family and the application year of all the distinct cited patent families	*DII* (List of *Patents Cited by Inventor / Examiner*)
Recent Technological Activity	The sum of the patent families the firm had applied for in the prior five years, for each firm-year observation	*DII*
Internal Focus	The average ratio of the citations that each patent family makes to prior patent families held by the focal firm to its total backward citations	*DII* (List of *Patents Cited by Inventor / Examiner*)
R&D Investment	Yearly R&D expenditure	*Scoreboard*
Profitability	Yearly *operating* profit	*Scoreboard*
Firm Size	Number of firm’s employees	*Scoreboard*

Variable	*Measure*	*Source*
Number of Claims	The average number of claims of each patent family, for each firm-year observation	*DII* (*Abstract*)
Number of Inventors	The average number of inventors of each patent family, for each firm-year observation	*DII* (*Inventors*)
Geographical Coverage – Patent Family Size	The average number of patents of each patent family, for each firm-year observation	*DII* (List of *Patent Number*)
Geographical Coverage – Designated States	The average number of designated states of each patent family, for each firm-year observation	*DII* (*Designated States*)
Number of Patents	The average number of patent families, for each firm-year observation	*DII*
Number of Granted Patents	The average number of patent families that have at least one granted patent as their member, for each firm-year observation	*DII* (*Patents Details* → *Patent Number*)
Year	Year dummies	
Industry	Industry dummies	*Scoreboard*
Country	Country dummies	*Scoreboard*

In *Appendix* D, we provide a description of the procedure that we follow in order to deploy each patent-based variable from its raw form (as given by *DII* and stored in our database). In particular, we present in a *SQL Query Style* how each patent-based variable has been built from the data that are stored in the tables of our MySQL database.

3.4 Statistical Analysis

The data consist of a panel of observations on firm-years. Panel or longitudinal data is a set of data which have both a cross-sectional and a time series dimension (Wooldridge, 2002). In order to construct the panel, we followed the same

139 firms for 7 years (i.e. balanced panel data), building a set of data of 973 firm-year observations. We choose balanced panel data instead of unbalanced in order to avoid the possible correlation of the idiosyncratic errors with the reason of the missing data, which can cause biased estimators (Wooldridge, 2002). The key issue with an unbalanced panel is determining why the panel is unbalanced. If the reason we have missing data is not correlated with the idiosyncratic errors, then the unbalanced panel causes no problems. However, if the reason is correlated with the idiosyncratic error, then this reason is considered as an unobserved factor that changes over time and affects the dependent variable, and thus, it can cause biased estimators. But since we had enough firms to construct an adequate sample size based on balanced panel data (139 firms), we chose not to run the risk of unbalanced panels.

With regard to the magnitude of impact, the empirical model that we estimate takes, in general, the following form: *Magnitude of Impact* = f (*Scientific Basis, Technological Collaboration, Technological Diversity, New-to-Firm Technological Domains, Unexplored Technological Areas, Recency of Technological Inputs, Recent Technological Activity, Internal Focus, R&D Investment, Firm Size, Controls)*. Since our dependent variable, firm's impact on technological evolution as represented by patent citations count, is a discrete, non-negative variable, conventional linear regression models are inappropriate (Fosfuri 2006). The reason is that the linear regression models assume homoskedastic, normally distributed errors, an assumption which is violated with count variables (Ahuja 2000, Operti and Carnabuci 2014).

The simplest model to handle count data is the Poisson regression model. Poisson models are based on the assumption that the dependent variable is drawn from a Poisson distribution, which has equal mean and variance (Penner-Hahn and Shaver 2005). However, if this assumption does not hold and the data exhibit overdispersion, Poisson models can lead to biased estimates in some contexts, especially with cross-sectional data (Somaya et al. 2007). A remedy for this is the negative binomial distribution, which is a generalization of the Poisson model. Negative binomial distribution is a two-parameter model that estimates an overdispersion parameter and produces correct standard errors for count data that are overdispersed (Gittelman and Kogut 2003) by incorporating an individual unobserved effect into the conditional mean (Operti and Carnabuci 2014). In our case, our dependent variable has a variance (*value: 146776.3*) that is larger than its mean (*value: 148.42*), suggesting the presence of overdispersion. Therefore, negative binomial distribution is the suitable model for our data.

Theoretically, to account for the unobserved heterogeneity of the negative binomial regression, that is, the possibility that firms differ on unmeasured characteristics, either fixed or random-effects specification can be used (Rothaermel and Hess 2007). Fixed-effects specification serves as a control for unobserved, ***time-invariant*** explanatory variables that may affect the response variable (Dushnitsky and Lenox 2005). Random-effects specification accounts also for unobserved heterogeneity, and, moreover, allows to include ***time-invariant*** explanatory variables, such as the type of industry and national origin (Vanhaverbeke et al. 2009). However, random-effects specification is based on the assumption that the time-invariant, unobserved effect is uncorrelated with all the time-variant explanatory variables. In other words, if the unobserved, time-invariant effect is correlated with some explanatory variables, the random-effects specification is inappropriate because the estimators it produces are inconsistent (Wooldridge, 2002). For example, if the random-effects specification model includes the firm's country of origin, which is a time-invariant variable that affects both patent forward citations (the response variable) and the R&D investment (an explanatory variable), then, according to Wooldridge (2002), the estimators of the model would be inconsistent.

In order to test this assumption, a Hausman specification test (Hausman 1978) must be applied. The Hausman specification test compares the fixed-effects' estimators, which are known to be consistent, with the random-effects' estimator that are efficient under the assumption being tested (i.e. the time-invariant, unobserved effect is uncorrelated with all the time-variant explanatory variables). If the random-effects' estimators are consistent, there should be no systematic difference between the two set of estimators. On the contrary, if the assumption is wrong, this will be reflected in a difference between the two set of estimators. The bigger the difference (the less similar are the two sets of estimators), the bigger the Hausman statistic. A large and significant Hausman statistic means a large and significant difference, and so the null hypothesis that both methods produce consistent estimators is rejected. In this case, you have reason to doubt the assumption on which random-effects specification is based and so, this test cannot be applied.

We build three models to test our hypotheses. The first model, which is our general model, includes all the firm-year observations of our sample. The Hausman specification test was significant in this model, indicating that the unobserved, time-invariant effect is correlated with some of our independent variables and that it is not safe to use a random-effects specification (chi2=89.83, Prob>chi2 = 0.0000). Hence, we estimated a panel fixed-effects negative binomial model.

In the second model, only the firm-year observations with patent citations count larger than zero were taken in account. We restrict the sample only to these firm-year patent families that have at least a minimum impact on technological evolution, in order to test the impact of the explanatory variables on the response variable without the effect of the non-influential firm-year patent families (e.g. patent families of a given year that have received zero citations). For this model, the Hausman specification test rejected the unbiasedness of the random-effect estimator, suggesting fixed-effects negative binomial model as the appropriate model (chi2=101.17, Prob>chi2 = 0.0000).

Finally, in the third model, to ensure that our findings were robust, we restrict the sample to observations in the top 75th percentile of the dependent variable (i.e. patent citations count≥4)[30]. We did this because we want to test if this restricted sample will lead to the same results as the results of the second model (Yayavaram and Ahuja 2008). The more similar the results, the higher the probability the estimated coefficients to be independent from the sample. The Hausman test of this model was again significant, and hence, the random-effects specification was rejected and the fixed-effect was applied (chi2=65.67, Prob>chi2 = 0.0004).

---

[30] A set of summary statistics for the dependent variable *Magnitude of Impact*

Magnitude of Impact (i.e. Patent Citations Count)				
	Percentiles	**Smallest**		
1%	0	0		
5%	0	0		
10%	0	0	**Obs**	973
25%	3	0	**Sum of Wgt.**	973
50%	23		**Mean**	1.484.214
		Largest	**Std. Dev.**	383.114
75%	132	2697		
90%	398	3095	**Variance**	146776.3
95%	697	3433	**Skewness**	7.495.121
99%	1706	6427	**Kurtosis**	9.217.691

We ran these models using the *xtnbreg* command in the *STATA* statistical software, with the fixed-effects option (*STATA 11)*. Moreover, for the estimated coefficients to be comparable, we standardized all independent variables before entering them into the regression models (Rothaermel and Hess 2007, Operti and Carnabuci 2014).

At this point, we have to note that, because *xtnbreg* uses conditional log likelihood, the dependent variable is logged[31]. And, consequently, the estimated coefficients of the negative binomial model represent semielasticities (i.e. the proportionate change in the conditional mean caused by a one-unit change in the explanatory variable) (Hoetker and Agarwal 2007). For example, if the estimated coefficient of the independent variable *Technological Collaboration* is 1.23, this means that the increase in *Technological Collaboration* by one unit will cause an increase of 1.23% in the dependent variable, holding other independent variables fixed. For example, if each patent family of a given firm had *X* co-assignees on average in year *t* and *X+1* in year *t+1*, and the aggregate number of citations received by its patents in year *t* is *Y*, then, in year *t+1*, the aggregate number of citations received would be *Y+Y*1.23%*. However, when the dependent variable is expressed in logs (e.g. *Recent Technological Activity*), the estimated coefficients of the negative binomial model represent elasticities, that is to say, the percentage change in the dependent variable given a change of one percentage point in the independent variable.

At this point, it is necessary to emphasize that although one would expect our models, as fixed-effects models, not to include time-invariant variables[32], we

[31] Apparently, this is the reason for not using logs in our dependent variable (e.g. patent citations count), although it is probably a large number.

[32] Fixed effects models are designed to study the causes of changes within a firm (or entity in general), and so they can control for unobserved time-invariant variables by introducing an additional parameter

nevertheless included them (e.g. industry and country dummies). The reason is that the method upon which the *STATA* command *xtnbreg* is build does not qualify as a true fixed-effects method because it does not accomplished what is usually desired in a such a method, namely, the control of all the time-invariant explanatory variables.

More specifically, the *xtnbreg* is based upon the work of Hausman et al. (1984), who proposed a conditional likelihood method for negative binomial regression. However, as explained in Allison and Waterman (2002), the problem with this method is that it allows for individual specific variation in the dispersion parameter rather than in the conditional mean. As a result, unlike other conditional likelihood methods (e.g. Logistic and Poisson fixed effects models)[33], you can put time-invariant covariates into *xtnbreg* model and get nonzero coefficient estimates for those variables[34].

Therefore, we include the time-invariant industry and country dummies in our models even if we used the fixed-effects specifications (Biotechnology serves as the reference industry and EU serves as the reference country). Time-invariant variables in fixed effects negative binomial models where the dependent variable was the number of patent citations were also used by George et al. (2008), Operti and Carnabuci (2014) and Miller et al. (2007)[35].

With regard to the breadth of technological impact, generally, our empirical model for the breadth of technological impact takes the following form: *Breadth of Impact* = f (*Scientific Basis*, *Technological Collaboration*, *Technological Diversity*,

---

for each individual in the sample. Depending on the type of the response variable, fixed effects models come in many forms: linear models for quantitative outcomes, logistic models for dichotomous outcomes, and Poisson regression models for count data (Allison 2009).

[33] In conditional likelihood, for each individual, the contribution to the likelihood function is conditioned on the sum of the repeated measures, thereby eliminating the individual specific parameters from the likelihood function (Allison 2009).

[34] The same conclusions were reached also by Guimarães (2008) and Greene (2007).

[35] All the three studies used STATA to estimate their equations.

New-to-Firm Technological Domains, *Unexplored Technological Areas, Recency of Technological Inputs, Recent Technological Activity*, *Internal Focus, R&D Investment, Firm Size, Controls).* As we mentioned in the section defining the dependent variables, we use three different measures of the breadth of technological impact.

The first two measures, *Beyond-Industry Impact* (*Model 4*) and *Beyond-Firm Impact* (*Model 5*), use patent citations count as dependent variable. As in the case of technological impact (*Models 1-3*), since the dependent variable is a discrete, non-negative variable and its variance is larger than its mean, we applied a negative binomial distribution model. Moreover, the Hausman specification test suggested that the random-effects specification will produce inconsistent estimators and, hence, the fixed-effects specification was chosen (chi2=54.59, Prob>chi2 = 0.0039). As before, we ran these models using the *xtnbreg* command with the fixed-effects option including dummies for the industry and the country effect.

The third metric of breadth of impact (*Models 6-8*), which is actually the distribution (or the concentration) of the technological classes that are affected by a firm's inventions, is a Herfindahl type measure, and, therefore a percentage. Based on previous studies, we found three different ways to handle models whose dependent variable is a percentage. For robustness, we test all of the three.

First, as Alcacer et al. (2009) suggested, one can use a generalized linear model (glm) with a logit link utilizing binomial family. In this approach, the dependent variable is treated as a binary response (*Model 6*). We ran this model using the *xtgee family(binomial) ling(logit)* command in *STATA*. This command fits population-averaged logit models. In these models, the time-invariant independent variables (e.g. industry and country) must be included since the population-averaged logit models do not control for unobserved, time-constant factors that affect the response variable.

Second, in *model 7*, we used a tobit specification, since Herfindahl measures are truncated at a lower limit of zero and an upper limit of one (Mowery and Ziedonis 2002). In this case, we used the *xttobit* command. *Xttobit* fits only random-effects tobit models since there is no command for a parametric conditional fixed-effects model, as there does not exist a sufficient statistic allowing the fixed effects to be conditioned out of the likelihood[36] (Banerjee and Cole 2010, Valentini and Di Guardo 2012). Consequently, industry and country dummies were included in this model.

Finally, in *model 8*, following Benner and Tushman (2002), we applied a simple linear regression (e.g. Ordinary Least Squares) with a fixed-effects specification as the Hausman test suggested (chi2=58.20, Prob>chi2 = 0.0003). The *xtreg* command was used for this model. *Xtreg* model is a true fixed-effect model since it controls for all the time-invariant variables. Therefore, we did not include dummies for industry and country.

As we mentioned earlier, because Herfindahl type measures suffer from bias due to the properties of small numbers, we consider only the patents that receive over three citations (Argyres and Silverman 2004). Furthermore, as before, we standardized all independent variables before entering them into the regression models. In addition, we have to note that models 6, 7, and 8 are all level-level models, meaning that the estimated coefficients are interpreted in the following way: the change in the dependent variable given a one-unit change in the independent variable produces. But when the independent variables is expressed in logs (e.g. *Recent Technological Activity*), the estimated coefficients represent the change in the

[36] See the command *help xttobit* in *STATA 11*

response variable caused by a change of one percentage point in the explanatory variable.

Moreover, it is important to note that, according to Wooldridge (2002), the estimated coefficients of the logit models (*model 6*) and the tobit models (*model 7*) are hard to be compared with those of the linear regression models (*model 8)*. So, in the Results section (i.e. *Chapter 4*), we will proceed in the comparison of the estimated coefficients within each model and not across these tree models.

3.5 Descriptive Statistics

Table 3.3 presents the descriptive statistics and the correlation values for all variables. In order to assess potential problems of multicollinearity, we calculated a more advanced measure of it, the VIF (variance inflation factor), for all the variables (*Appendix E*).

The only variables that seem to be suffering from multicollinearity are the *Recent Technological Activity* and the *Number of Patents.* The value of the correlation of these variables is 0.95 and the value of the VIF for *Recent Technological Activity* is 11.81 and for *Number of Patents* is 15,39. As a rule of thumb, the critical value of VIF, above which the econometric models consider to suffer from multicollinearity is the value of 10 (Operti and Carnabuci 2014). The fact that these VIF values are above the recommended cutoff value can cause misleading results in the regression analysis regarding these specific independent variables, since the standard errors in cases of multicollinearity are increased (Wooldridge, 2002). Therefore, we must be cautious in our interpretation of the result of these specific variables.[37]

[37] However, we did not exclude any of these variables from our models because both of them are relevant variables and important predictors of our response variables. If we omit these important variables, we will cause the “omitted variable bias” which is a form of misspecification analysis.

With regard to the rest variables, according to *Appendix E*, the mean value of VIF is 3.47, while the highest VIF value is 5.23 (*Firm Size*), way below the critical value of 10. Thus, for all the rest variables, multicollinearity is not an issue in our data.

Table 3.3 Descriptive Statistics

Variable	Mean	Std dev	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. Impact On Technological Evolution	148.42	383.11	0	6427	1.00																						
2. Breadth of Impact – *Beyond-Industry Impact*	0.60	2.10	0	30	0.65	1.00																					
3. Breadth of Impact – *Beyond-Firm Impact*	8.31	12.77	0	83	0.70	0.48	1.00																				
4. Breadth of Impact – *Herfindahl Type Breadth*	0.57	0.34	0	0.97	0.24	0.13	0.48	1.00																			
5. Scientific Basis – *Non Patent References*	4.73	8.54	0	106.57	0.03	0.23	0.03	0.01	1.00																		
6. Scientific Basis – *Scientific Inventions*	0.01	0.06	0	1	-0.05	-0.03	-0.08	-0.08	0.00	1.00																	
7. Technological Collaboration	2.30	1.41	0	12	0.15	0.20	0.14	0.07	0.29	-0.03	1.00																

Variable	Mean	Std dev	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
8. Technological Diversity – *IPC Diversity*	2.71	0.94	0	11	0.01	0.07	0.02	0.11	0.28	-0.03	0.40	1.00															
9. Technological Diversity – *Originality*	0.73	0.24	0	0.97	0.18	0.06	0.35	0.63	-0.07	-0.11	-0.05	0.03	1.00														
10. New-to-Firm Technological Domains	6.38	6.63	0	38	0.46	0.22	0.58	0.55	-0.18	-0.08	-0.03	-0.02	0.54	1.00													
11. Unexplored Technological Areas	5.73	6.15	0	63.77	0.28	0.30	0.35	0.15	0.60	-0.09	0.39	0.19	0.08	0.12	1.00												
12. Recency of Technological Inputs	8.38	3.45	0	30.25	-0.01	-0.04	0.06	0.08	-0.09	-0.09	-0.05	-0.03	0.33	0.25	0.16	1.00											
*13. Recent Technological Activity**	5.51	1.78	0	9.10	0.39	0.17	0.45	0.64	-0.17	-0.02	-0.20	-0.13	0.68	0.64	-0.04	0.09	1.00										
14. Internal Focus	0.19	0.13	0	0.83	0.01	-0.02	-0.01	0.12	-0.05	0.00	-0.17	0.02	0.16	0.02	-0.20	-0.17	0.30	1.00									
*15. R&D Investment**	4.75	1.55	1.54	8.76	0.41	0.28	0.29	0.34	-0.01	-0.04	0.03	-0.01	0.35	0.38	0.09	-0.02	0.62	0.16	1.00								
*16. Profitability**	12.35	0.56	0.02	14.16	0.17	0.22	0.11	0.14	-0.03	-0.04	0.05	-0.01	0.13	0.20	0.08	0.02	0.27	0.02	0.48	1.00							
*17. Firm Size**	8.41	1.88	2.70	11.71	0.37	0.16	0.38	0.43	-0.26	-0.07	-0.11	-0.18	0.50	0.60	0.04	0.28	0.68	-0.02	0.74	0.40	1.00						

Variable	Mean	Std dev	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
18. Number of Claims	2.21	1.50	0	16.10	-0.06	0.04	-0.10	-0.11	0.34	0.10	0.18	0.41	-0.21	-0.23	0.13	-0.21	-0.24	-0.03	-0.11	-0.06	-0.36	1.00					
19. Number of Inventors	4.70	2.90	0	39.50	-0.01	0.06	-0.08	-0.09	0.22	-0.04	0.43	0.34	-0.14	-0.19	0.16	-0.13	-0.19	0.01	0.11	0.05	-0.17	0.25	1.00				
20. Geographical Coverage – *Patent Family Size*	4.95	3.08	0	19	0.17	0.21	0.17	0.17	0.37	-0.13	0.58	0.52	-0.03	-0.02	0.38	0.01	-0.21	-0.07	0.09	0.07	-0.07	0.20	0.45	1.00			
21. Geographical Coverage – *Designated States **	4.09	1.07	0	4.86	-0.09	0.03	-0.19	-0.20	0.21	0.02	0.31	0.34	-0.13	-0.27	0.21	0.10	-0.35	-0.01	0.08	0.08	-0.16	0.25	0.40	0.47	1.00		
*22. Number of Patents**	3.73	1.82	0	7.28	0.42	0.20	0.50	0.69	-0.17	-0.06	-0.15	-0.13	0.72	0.71	-0.01	0.13	0.95	0.25	0.62	0.28	0.72	-0.27	-0.19	-0.17	-0.33	1.00	
23. Number of Granted Patents	0.23	0.23	0	1	0.32	0.26	0.53	0.42	0.29	-0.14	0.40	0.22	0.21	0.23	0.49	0.08	0.06	-0.14	0.05	0.03	0.12	0.05	0.04	0.57	0.01	0.12	1.00
*Variables are logged.																											

RESULTS

4 RESULTS

4.1 Results

Concerning the magnitude of impact, as we mentioned earlier in subsection *3.4*, we build three models to test our hypotheses regarding the firms' impact on technological evolution (*Models 1-3*). In the first model (*Model 1*), we included all the firm-year observations of our sample. The second model *(Model 2)* includes only these firm-year patent families that have at least a minimum effect on subsequent patent families (i.e. patent citations count≥1). We build this model in order to test our hypotheses without the effect of the non-influential firm-year patent families. Finally, in the third model *(Model 3)*, to ensure that our findings were robust, we restrict the sample to observations in the top 75th percentile of the dependent variable (i.e. patent citations count≥4). The more similar the estimators of the restricted sample's model with the estimators of the previous models, the higher the probability the estimated coefficients to be independent from the sample. In all three models, the Hausman specification test was significant, indicating that it is not safe to use a random-effects specification. Instead, we estimated a panel fixed-effects negative binomial model. As far as the goodness-of-fit statistics are concerned, we used the value of the log likelihood in order to compare the goodness-of-fit of the first three models. As we know, the goal of a model is to find values for the coefficients that maximize the value of the likelihood function, that is, to find the set of parameter estimates that make the data most likely. The closer to zero the value of the log likelihood, the better the fit of the model. Among the models 1-3, model 3 shows the best goodness-of-fit, since its log likelihood is smaller than models 1 and 2 (*Model1*: -1966.90, *Model2*: -1915.80, *Model3*: -1674.12)

With regard to breadth of impact, as was noted previously, we use three distinct measures of the breadth of technological impact in order to capture three different dimensions of the breadth of impact. The *Beyond-Industry Impact* (*Model 4*), which focuses on the firm's impact on industries other than its own, the *Beyond-Firm Impact* (*Model 5*), which focuses on the impact on technological areas different from the areas that the focal firm operates in, and, finally, the *Herfindahl Type Breadth* (*Model 5-8*), which focuses on the concentration of the international patent classes of the affected patent families. Thus, the coefficients of a given independent variable estimate the impact of this variable on the three different aspects of breadth of impact.

For the *Beyond-Industry Impact* and the *Beyond-Firm Impact*, we applied a fixed-effects negative binomial model. For the *Herfindahl Type Breadth*, which is actually a percentage, we used three different models. The first model is a generalized linear model (glm) with a logit link utilizing binomial family (*Model 5*), the second is a tobit specification (the dependent variable is truncated at a lower limit of zero and an upper limit of one) and the third model is a linear regression model with a fixed-effects specification. As far as the fit statistics of the *Models 4-8* are concerned, the value of the log likelihood can be used in models 4, 5, and 7 (*Model 4*: -171.19, *Model 5*: -968.89, *Model 7*: -700.74). In *Model 6*, the value of the Deviance is offered, which is another method for the assessment of the goodness-of-fit. Deviance compares the values predicted by the fitted model and those predicted by "the most complete model we could fit". Evidence for model lack-of-fit occurs when the value of Deviance is large (Deviance in

Model 6: 28.37). Finally, in *Model 8*, *STATA* provides the value of the adjusted R-squared. The value of the adjusted R-squared in *Model 8* is 0.21, meaning that the model explains 21 percent of the variability of the response variable.

However, we have to note that there exist two important differences between *Models 4-5* and *Models 6-8*. First, in *Models 6-8* there is not any lagged value, and consequently the observations are substantially more than in models *Models 4-5*, since the observations of the first year of our sample (i.e. 2003) are included in *Models 6-8* but not in *Models 4-5*. We did not include any lagged value in these models because the lagged values of the variables *R&D Investment* and *Profitability*, which we used in *Models 1-5*, were insignificant in *Models 6-8*. So, we excluded them in order to gain in observations. Second, we included the control variable *Number of Citations Received* only in *Models 6-8* where the dependent variable is the *Herfindahl Type Breadth* and not in any other model.

Table 4.1 Panel Fixed-effects Negative Binomial Models Explaining the Factors that Affect the Firm's Ability to Shape Technological Evolution and to Exert a Broad Influence on Technological Change

Variables	**Model 1** **Magnitude of Impact** *(All Observations)*	**Model 2** **Magnitude of Impact** *(Patent Citations Count>0)*	**Model 3** **Magnitude of Impact** *(Patent Citations Count ≥4 ––– 75th Percentile)*	**Model 4** **Breadth of Impact – *Beyond-Industry Impact*** *(Fixed-effects Negative Binomial)*	**Model 5** **Breadth of Impact – *Beyond-Firm Impact*** *(Fixed-effects Negative Binomial)*	**Model 6** **Breadth of Impact – *Herfindahl Type Breadth I*** *(Glm, Logit Link, Binomial Family)*	**Model 7** **Breadth of Impact – *Herfindahl Type Breadth II*** *(Tobit Specification)*	**Model 8** **Breadth of Impact – *Herfindahl Type Breadth III*** *(Fixed-effects Linear Regression)*
Scientific Basis – *Non Patent References*	0.006 (0.028)	0.009 (0.027)	0.002 (0.029)	0.128 (0.163)	0.077 (0.064)	-0.030 (0.034)	-0.004 (0.005)	-0.008 (0.006)
Scientific Basis – *Scientific Inventions*	-0.006 (0.038)	-0.005 (0.039)	0.016 (0.044)	0.559+ (0.315)	0.155+ (0.085)	0.129*** (0.037)	0.028*** (0.008)	0.029** (0.009)
Technological Collaboration	-0.031 (0.040)	-0.029 (0.040)	-0.024 (0.045)	0.339+ (0.197)	0.081 (0.067)	-0.038 (0.033)	-0.006 (0.006)	-0.009 (0.007)

Variables	**Model 1** Magnitude of Impact *(All Observations)*	**Model 2** Magnitude of Impact *(Patent Citations Count>0)*	**Model 3** Magnitude of Impact *(Patent Citations Count ≥4 ––– 75th Percentile)*	**Model 4** Breadth of Impact – ***Beyond-Industry Impact*** *(Fixed-effects Negative Binomial)*	**Model 5** Breadth of Impact – ***Beyond-Firm Impact*** *(Fixed-effects Negative Binomial)*	**Model 6** Breadth of Impact – ***Herfindahl Type Breadth I*** *(Glm, Logit Link, Binomial Family)*	**Model 7** Breadth of Impact – ***Herfindahl Type Breadth II*** *(Tobit Specification)*	**Model 8** Breadth of Impact – ***Herfindahl Type Breadth III*** *(Fixed-effects Linear Regression)*
Technological Collaboration ***Squared***	0.000 (0.014)	-0.002 (0.013)	-0.004 (0.017)					
Technological Diversity – *IPC Diversity*	0.067 (0.044)	0. 065 (0.044)	0.061 (0.051)	-0.987* (0.395)	-0.133 (0.098)	0.152** (0.055)	0.028*** (0.008)	0.019+ (0.011)
Technological Diversity – *Originality*	-0.245**** (0.051)	-0.275**** (0.051)	-0.298**** (0.052)	0.254 (0.347)	0.154 (0.131)	0.533**** (0.078)	0.123**** (0.012)	0.102**** (0.015)
New-to-Firm Technological Domains	0.003 (0.025)	0.005 (0.025)	0.006 (0.025)	0.094 (0.116)	-0.021 (0.035)	0.036 (0.038)	0.004 (0.005)	0.004 (0.006)

Variables	**Model 1** Magnitude of Impact *(All Observations)*	**Model 2** Magnitude of Impact *(Patent Citations Count>0)*	**Model 3** Magnitude of Impact *(Patent Citations Count ≥4 ––– 75th Percentile)*	**Model 4** Breadth of Impact – *Beyond-Industry Impact* *(Fixed-effects Negative Binomial)*	**Model 5** Breadth of Impact – *Beyond-Firm Impact* *(Fixed-effects Negative Binomial)*	**Model 6** Breadth of Impact – *Herfindahl Type Breadth I* *(Glm, Logit Link, Binomial Family)*	**Model 7** Breadth of Impact – *Herfindahl Type Breadth II* *(Tobit Specification)*	**Model 8** Breadth of Impact – *Herfindahl Type Breadth III* *(Fixed-effects Linear Regression)*
*New-to-Firm Technological Domains **Squared***	-0.007 (0.007)	-0.005 (0.007)	-0.005 (0.007)					
Unexplored Technological Areas	-0.074*** (0.019)	-0.071*** (0.019)	-0.070** (0.020)	-0.151 (0.165)	-0.037 (0.047)	0.047 (0.038)	0.007 (0.005)	0.006 (0.006)
Recency of Technological Inputs	0.059 (0.036)	0.080* (0.037)	0.089* (0.039)	-0.321 (0.307)	-0.021 (0.080)	-0.006 (0.047)	-0.002 (0.007)	-0.003 (0.010)
Recent Technological Activity	-0.127 (0.118)	-0.143 (0.118)	-0.096 (0.130)	0.828 (1.509)	0.022 (0.251)	0.241*** (0.069)	0.038** (0.015)	-0.008 (0.026)
Internal Focus	0.042	0.009	-0.013	0.262	-0.001	0.023	0.003	-0.004

Variables	**Model 1** Magnitude of Impact *(All Observations)*	**Model 2** Magnitude of Impact *(Patent Citations Count>0)*	**Model 3** Magnitude of Impact *(Patent Citations Count ≥4 ––– 75th Percentile)*	**Model 4** Breadth of Impact – *Beyond-Industry Impact* *(Fixed-effects Negative Binomial)*	**Model 5** Breadth of Impact – *Beyond-Firm Impact* *(Fixed-effects Negative Binomial)*	**Model 6** Breadth of Impact – *Herfindahl Type Breadth I* *(Glm, Logit Link, Binomial Family)*	**Model 7** Breadth of Impact – *Herfindahl Type Breadth II* *(Tobit Specification)*	**Model 8** Breadth of Impact – *Herfindahl Type Breadth III* *(Fixed-effects Linear Regression)*
	(0.030)	(0.031)	(0.034)	(0.269)	(0.070)	(0.038)	(0.005)	(0.007)
R&D Investment	-0.076 (0.102)	-0.065 (0.101)	-0.100 (0.108)	1.202 (1.174)	0.255 (0.238)	0.033 (0.077)	-0.004 (0.012)	-0.027 (0.027)
Lagged R&D Investment	0.148 (0.093)	0.112 (0.092)	0.172+ (0.097)	1.687 (1.036)	0.029 (0.226)			
Firm Size	-0.234 (0.144)	-0.225 (0.146)	-0.172 (0.155)	-2.503 (1.670)	-0.215 (0.274)	-0.007 (0.080)	0.003 (0.013)	0.086+ (0.045)
Number of Patents	1.747**** (0.075)	1.669**** (0.075)	1.687**** (0.083)	1.341* (0.627)	1.352**** (0.181)	0.284+ (0.148)	0.076*** (0.021)	0.046+ (0.027)

Variables	**Model 1** Magnitude of Impact *(All Observations)*	**Model 2** Magnitude of Impact *(Patent Citations Count>0)*	**Model 3** Magnitude of Impact *(Patent Citations Count ≥4 ––– 75th Percentile)*	**Model 4** Breadth of Impact – *Beyond-Industry Impact* *(Fixed-effects Negative Binomial)*	**Model 5** Breadth of Impact – *Beyond-Firm Impact* *(Fixed-effects Negative Binomial)*	**Model 6** Breadth of Impact – *Herfindahl Type Breadth I* *(Glm, Logit Link, Binomial Family)*	**Model 7** Breadth of Impact – *Herfindahl Type Breadth II* *(Tobit Specification)*	**Model 8** Breadth of Impact – *Herfindahl Type Breadth III* *(Fixed-effects Linear Regression)*
Number of Claims	0.009 (0.036)	0.024 (0.036)	0.019 (0.040)	0.290 (0.408)	0.102 (0.079)	0.017 (0.040)	0.003 (0.005)	0.005 (0.007)
Number of Inventors	0.087** (0.033)	0.083* (0.032)	0.091* (0.034)	0.061 (0.170)	0.067 (0.078)	0.080+ (0.044)	0.015** (0.005)	0.020** (0.007)
Geographical Coverage – *Patent Family Size*	0.082+ (0.044)	0.075+ (0.044)	0.110* (0.048)	0.461 (0.406)	0.177+ (0.106)	-0.086* (0.048)	-0.011 (0.008)	-0.010 (0.010)
Geographical Coverage – *Designated States*	0.235*** (0.068)	0.203** (0.070)	0.188** (0.072)	-0.960* (0.406)	0.132 (0.174)	-0.189* (0.093)	-0.016 (0.013)	-0.012 (0.021)

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Magnitude of Impact *(All Observations)*	**Magnitude of Impact** *(Patent Citations Count>0)*	**Magnitude of Impact** *(Patent Citations Count ≥4 —– 75th Percentile)*	**Breadth of Impact – *Beyond-Industry Impact*** *(Fixed-effects Negative Binomial)*	**Breadth of Impact – *Beyond-Firm Impact*** *(Fixed-effects Negative Binomial)*	**Breadth of Impact – *Herfindahl Type Breadth I*** *(Glm, Logit Link, Binomial Family)*	**Breadth of Impact – *Herfindahl Type Breadth II*** *(Tobit Specification)*	**Breadth of Impact – *Herfindahl Type Breadth III*** *(Fixed-effects Linear Regression)*
Number of Granted Patents	0.108** (0.037)	0.101** (0.036)	0.112*** (0.040)	-0.031 (0.346)	0.002 (0.082)	0.063 (0.053)	0.013+ (0.007)	0.017+ (0.009)
Profitability	0.132** (0.045)	0.127** (0.045)	0.148** (0.046)	-0.232 (0.309)	-0.086 (0.114)	0.011 (0.018)	0.002 (0.003)	0.004 (0.004)
Lagged Profitability	0.009* (0.003)	0.008* (0.004)	0.008* (0.004)	0.013 (0.063)	0.026* (0.011)			
Number of Citations Received						-0.197 (0.132)	-0.053** (0.017)	-0.044* (0.021)
Year 2003						1.738**** (0.268)	0.315**** (0.039)	0.296**** (0.048)

Variables	**Model 1** Magnitude of Impact *(All Observations)*	**Model 2** Magnitude of Impact *(Patent Citations Count>0)*	**Model 3** Magnitude of Impact *(Patent Citations Count ≥4 ––– 75th Percentile)*	**Model 4** Breadth of Impact – ***Beyond-Industry Impact*** *(Fixed-effects Negative Binomial)*	**Model 5** Breadth of Impact – ***Beyond-Firm Impact*** *(Fixed-effects Negative Binomial)*	**Model 6** Breadth of Impact – ***Herfindahl Type Breadth I*** *(Glm, Logit Link, Binomial Family)*	**Model 7** Breadth of Impact – ***Herfindahl Type Breadth II*** *(Tobit Specification)*	**Model 8** Breadth of Impact – ***Herfindahl Type Breadth III*** *(Fixed-effects Linear Regression)*
Year 2004	3.755**** (0.120)	3.600**** (0.118)	3.469**** (0.138)	20.029 (1925.515)	4.487**** (0.349)	1.556**** (0.256)	0.288**** (0.036)	0.268**** (0.044)
Year 2005	3.419**** (0.114)	3.269**** (0.112)	3.139**** (0.132)	19.681 (1925.515)	4.221**** (0.337)	1.490**** (0.228)	0.277**** (0.034)	0.260**** (0.040)
Year 2006	3.083**** (0.106)	2.929**** (0.104)	2.806**** (0.124)	19.201 (1925.515)	3.779**** (0.325)	1.265**** (0.215)	0.238**** (0.031)	0.222**** (0.037)
Year 2007	2.483**** (0.103)	2.346**** (0.101)	2.240**** (0.120)	18.813 (1925.515)	3.015**** (0.322)	1.170**** (0.188)	0.218**** (0.027)	0.204**** (0.031)
Year 2008	1.637**** (0.102)	1.493**** (0.100)	1.390**** (0.118)	17.515 (1925.515)	2.136**** (0.324)	0.747**** (0.145)	0.143**** (0.022)	0.139**** (0.024)

Variables	**Model 1** Magnitude of Impact *(All Observations)*	**Model 2** Magnitude of Impact *(Patent Citations Count>0)*	**Model 3** Magnitude of Impact *(Patent Citations Count ≥4 ––– 75th Percentile)*	**Model 4** Breadth of Impact – *Beyond-Industry Impact* *(Fixed-effects Negative Binomial)*	**Model 5** Breadth of Impact – *Beyond-Firm Impact* *(Fixed-effects Negative Binomial)*	**Model 6** Breadth of Impact – *Herfindahl Type Breadth I* *(Glm, Logit Link, Binomial Family)*	**Model 7** Breadth of Impact – *Herfindahl Type Breadth II* *(Tobit Specification)*	**Model 8** Breadth of Impact – *Herfindahl Type Breadth III* *(Fixed-effects Linear Regression)*
USA (EU omitted)	-0.264 (0.249)	-0.292 (0.254)	-0.100 (0.277)	-4.202* (1.792)	-0.753 (0.472)	-0.210 (0.132)	-0.044* (0.018)	
JAPAN (EU omitted)	0.095 (0.224)	-0.033 (0.224)	0.058 (0.244)	11.506 (1331.394)	0.049 (0.420)	0.087 (0.073)	0.020 (0.013)	
Pharmaceuticals (Biotechnology omitted)	0.250 (0.330)	0.155 (0.339)	0.347 (0.356)	-0.378 (2.412)	0.662 (0.588)	0.276* (0.130)	0.020 (0.022)	
Chemicals (Biotechnology omitted)	-0.090 (0.287)	-0.226 (0.295)	-0.065 (0.319)	-0.875 (1.733)	-0.239 (0.517)	-0.225** (0.077)	-0.046** (0.016)	
_cons	-0.419 (0.285)	0.010 (0.295)	-0.260 (0.320)	-17.273 (1925.515)	-2.322**** (0.591)	-0.296 (0.241)	0.469**** (0.033)	0.487**** (0.036)

Variables	**Model 1** Magnitude of Impact *(All Observations)*	**Model 2** Magnitude of Impact *(Patent Citations Count>0)*	**Model 3** Magnitude of Impact *(Patent Citations Count ≥4 —– 75th Percentile)*	**Model 4** Breadth of Impact – *Beyond-Industry Impact* *(Fixed-effects Negative Binomial)*	**Model 5** Breadth of Impact – *Beyond-Firm Impact* *(Fixed-effects Negative Binomial)*	**Model 6** Breadth of Impact – *Herfindahl Type Breadth I* *(Glm, Logit Link, Binomial Family)*	**Model 7** Breadth of Impact – *Herfindahl Type Breadth II* *(Tobit Specification)*	**Model 8** Breadth of Impact – *Herfindahl Type Breadth III* *(Fixed-effects Linear Regression)*
Log likelihood	-1966.90	-1915.80	-1674.12	-171.19	-968.89		700.74	
Deviance						28.37		
Adjusted R-squared								0.21
N	776	681	557	379	759	682	682	682

Standard errors in parentheses

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

Regarding *Hypothesis 1a*, the estimated coefficients of *Scientific Basis (i.e. Non Patent Reference* and *Scientific Inventions*) are insignificant in all models. These results are in accordance with Phene et al. (2006) who did not find any significant impact of the number of citations to sources such as academic journals, dissertations and theses (*i.e. Non Patent Reference*) on highly-cited inventions. Trying to interpret these findings, we refer to Sapsalis et al. (2006) who stated that scientific literature is a public information which is available to all inventors, and hence does not provide any specific additional value to the citing patent. Therefore, our results seem to reject *Hypothesis 1a* that the firm's scientific basis is positively related to its technological impact.

With respect to *Hypothesis 1b*, the results in *Table 4.1*, partly, but not fully, support this hypothesis. In particular, only the *Scientific Inventions* appears to exert an influence on technological breadth. More specifically, the estimated coefficients of *Scientific Inventions* are positive and significant in model 4 (*Beyond-Industry Impact*) and 5 (*Beyond-Firm Impact*) at $p < 0.1$, in models 6 and 7 at $p < 0.001$, and in model 8 at $p < 0.01$ (*Herfindahl Type Breadth*). Consequently, we find support for the notion that the more a firm relies on scientific knowledge to develop new technologies, the broader the firm's impact on technological evolution, on the condition that the scientific knowledge is measured as the number of patent families that have been applied by the firm in cooperation with a university (i.e. *Scientific Inventions*).

Contrary to our expectations, the coefficients of *Technological Collaboration* and *Technological Collaboration Squared* are insignificant in all models and thus the *Hypothesis 2a* is not supported. With respect to the variable *Technological Collaboration*, our results are in line with Srivastava and Gnyawali (2011) who also did not find any significant association between the number of patent assignees and the highly-cited inventions. However, we have to emphasize that these findings do not suggest that technological collaboration does not lead to influential inventions, but rather, that influential inventions can be generated with the same probability by unilateral R&D activities. Moreover, with respect to the variable *Technological Collaboration Squared*, the negative effect that we had hypothesized that it would be revealed beyond a certain value of *Technological Collaboration* did not appeared as well, in our results. Therefore, in general, we can state that the firm's impact on technological evolutions is unrelated to firm's technological collaboration.

Moreover, the same results were found for *Hypothesis 2b*. More specifically, the estimated coefficients of *Technological Collaboration* are insignificant in all models of breadth of impact, with the exception of model 5, in which *Technological Collaboration* appears to exert a positive influence on the *Beyond-Industry Impact*. However, because this impact is significant only at $p < 0.1$, no firm conclusion can be drawn about this specific relationship. Therefore, we can argue that the findings suggest that the extent to which a firm collaborates with other knowledge generating organizations to develop new inventions is not associated with the firm's breadth of impact.

With regard to the effect of *Technological Diversity* on technological impact, contrary to *Hypothesis 3a*, the coefficients of *Originality* are negative and statistically significant in all models, while the coefficients of *IPC Diversity* are all insignificant. We remind readers that we use two measures for *Technological Diversity*: *IPC Diversity* and *Originality*[38]. In particular, with respect to *Originality*, we can state that the more narrow the technological roots upon which an invention is built, the higher its influence on future inventions. It is probably, the high degree of specialization that results from the experimentation with certain but limited technological areas which make the inventions

[38] The first measure, the *IPC Diversification* is based on the average number of IPC of each patent family, while the second measure, the *Originality*, is a Herfindahl type measure which computes the dispersion of a patent family's backward citations across different IPC.

with lower technological diversity to be more influential. Another possible explanation for these results could be the high degree of complexity and difficulty that technologically diversified inventions face (Harhoff and Wagner 2006), a fact which may reduce their quality. Similar findings with the results regarding *IPC Diversity* have been reported by Nerkar (2003) who measured *Technological Diversity* as *IPC Diversity* and did not find any relationship between technological Diversity and technological impact.

In *Hypothesis 3b*, we proposed that the more technologically diversified a firm is, the wider the breadth of the firm's impact on technological evolution. However, our results provide mixed evidence on the relationship between the firm's technological diversity and the breadth of its impact. More specifically, the estimated coefficient of *IPC Diversity* is negative and significant in model 4 at $p < 0.05$. However, in models 6 and 7 the coefficients are positive and significant at $p < 0.01$ and at $p < 0.001$, respectively, while in model 8, they are positive and significant at $p < 0.1$. Additionally, the coefficients of *Originality* are also positive and statistically significant in models 6, 7, and 8 at $p < 0.001$. From these findings, we can conclude that the breadth of impact that pertains to a firm's technological impact outside of the boundaries of its industry is affected negatively by technologically diversity (as the coefficient of *IPC Diversity* in model 4 demonstrates). In others words, it is the narrowness of the technological domains and the possible specialization that derives from this narrowness that seems to positively influence firm's capability to develop inventions that encompass components of knowledge so profound and so general that can be useful to inventions originating in distant technological domains. On the other hand, when the breadth of impact is measured as a Herfindahl index, *Hypothesis 3b* is strongly supported and the *Technological Diversity* appears to be a significant positive predictor of the breadth of impact (as the coefficients of *IPC Diversity* and *Originality* in models 6, 7 and 8 indicate).

In *Hypothesis 4a,* we had proposed that as the experimentation with new-to-firm technologies increases, the impact of a firm's patent increases as well, but up to a point; beyond this point, the experimentation with new-to-firm technologies will affect negatively the firm's capability to develop influential inventions. Firstly, with regard to the hypothesized positive effect of *New-to-Firm Technological Domains* on technological evolution, the findings in *Table 4.1* don't indicate any such effect. In all models, the estimated coefficients of *New-to-Firm Technological Domains* are insignificant. These results can be explained if we consider the unfamiliarity with the new knowledge components that inventors face whenever they enter a new technological domain. Fleming (2001) emphasized this other side of the coin, by suggesting that the usage of familiar components improves inventors' ability to select the best components and to recombine them more successfully. In support of his argument, he presented evidence which indicated that inventors' knowledge familiarity is positively related to their inventions' usefulness. Consequently, the possible positive effect of experimenting with new technologies may be counteracted by the negative effect of unfamiliarity and, thus, the net effect of *New-to-Firm Technological Domains* on technological evolution becomes insignificant. Secondly, neither the negative effect of *New-to-Firm Technological Domains* beyond a certain threshold can be supported by our findings since the estimated coefficients of *New-to-Firm Technological Domains Squared* are insignificant in models 1-3.

Regarding *Hypothesis 4b*, the results in *Table 4.1* do not support it. In particular, the estimated coefficients of *New-to-Firm Technological Domains* are insignificant in all models. Therefore, we can conclude that the degree to which a firm enters new technological domains is irrelevant to the breadth of its impact on technological evolution.

Furthermore, our findings strongly reject *Hypothesis 5a*, since the estimated coefficients of *Unexplored Technological Areas* are negative and statically significant at $p < 0.01$ in all three first models (i.e. models of magnitude of impact). As mentioned before, this variable is computed as the average number of citations to prior patent families each patent family makes, for each firm-year observation. Thus, and after the aforementioned reversal of the sign of the variable, a high value indicates a lack of references to prior patent families which, in turn, probably denotes a firm whose R&D efforts focus on the development of inventions in technological areas that are largely unexplored. On the contrary, a low value of this variable indicates a high number of backward citations which, in turn, manifests – *presumably* – a firm whose R&D efforts rely on existing knowledge and on incremental technologies. Therefore, these negative and statically significant coefficients reveal that the more a firm develops patented inventions in technological areas where there exists a plethora of technological antecedents, as opposed to unexplored technological areas, the more the citations these patents receive and, thus, the greater the firm's impact on technological progress. However, these results did not take us by surprise since they are in accord with the empirical evidence of Sorenson and Fleming (2004) and Fleming (2001), who also found the number of backward citations (citations made) to be positively related to the number of forward citations (citations received).

Furthermore, the estimated coefficients of *Unexplored Technological Areas* are insignificant in models 4-8 (i.e. models of breadth of impact). These results are not contrary to the results reported by Banerjee and Cole (2010) who found that the total count of backward citations are neutral or slightly positive to the Herfindahl type breadth. Therefore, *Hypothesis 5b* is not supported and, consequently, we can come to the conclusion that the degree to which a firm experiments with unexplored technological areas has no effect on the breadth of its technological impact.

As far as the *Hypothesis 6a* is concerned, in general, the findings in *Table 4.1* support this hypothesis. Specifically, we have argued that the more recent the technologies that a firm incorporates into its inventions (as reflected in the backward citations of its patents), the greater the firm's impact on technological evolution. The results seem to support this hypothesis since, in models 2 and 3, the estimated coefficients of *Recency of Technological Inputs* are positive and significant at $p < 0.05$ (i.e. the more recent the inputs, the more the impact). However, the coefficient in model 1 is insignificant. Therefore, we can argue that the recency of technological inputs affects positively only the impact of the firm-year patent families that have played some role in the shaping of technological change, that is to say, in our case, the patent families of a given year of a given firm that have received at least one citation (Model 2) or at least 4 citations (Model 3). This effect, however, becomes insignificant when we include all the firm-year observations (Model 1). Consequently, we can argue that the more recent the technologies a firm incorporates into its inventions, the greater its impact on technological evolution, provided that we are talking about the firm-year patent families that have at least a minimum impact on the evolution of technology. More generally, we can propose that this specific effect is present only in those firms which invest relatively large amounts of resources in R&D every year (as all the firms of our sample) and their patented inventions on a yearly basis have at least a minimum impact on subsequent technological inventions. Or, in other words, this effect is significant only among influential firms and just among the firms that present high R&D expenditures.

In addition, our findings reject our expectations concerning *Hypothesis 6b*. The estimated coefficients of *Recency of Technological Inputs* are insignificant in models 4-8.

Thus, in general terms, the maturity of the technologies that a firm incorporates into its inventions does not seem to affect its technological breadth of impact.

Concerning *Hypothesis 7a*, findings do not suggest any relation between *Recent Technological Activity* and magnitude of impact. In all the first three models, the estimated coefficients are insignificant. Consequently, according to our results, the technological reputation that a firm gains by having developed a great number of new technological achievements in the past years does not appear to exert any influence on the number of citations that its current inventions receive. In addition, as we mentioned in subsection 3.5, the variable *Recent Technological Activity* is highly correlated with the control variable *Number of Patents*, a fact which affects the standard errors of these two variables. Thus, we decided to estimate again *Model 1-3* but without controlling for the *Number of Patents* in order to test the cumulative character of technological capabilities (i.e. the higher the *recent* technological activity, the higher the *current* technological activity, and, thus the greater the technological impact). In this case, the estimated coefficient of *Recent Technological Activity* was positive and significant at $p < 0.01$, at $p < 0.05$, and at 0.10 in models 1, 2 and 3 respectively and relatively very high (0.457, 0.348 and 0.325 in models 1, 2 and 3 respectively), supporting the cumulative character of technological capabilities, even if in our models this effect cannot be revealed.

Additionally, findings in *Table 4.1* partly, but not fully, support *Hypothesis 7b (Recent Technological Activity)*. More specifically, while the estimated coefficients in models 4 (*Beyond-Industry Impact*) and 5 (*Beyond-Firm Impact*) are insignificant, the coefficients in two of the three models of *Herfindahl Type Breadth* are positive and significant (in model 6 at $p < 0.001$ and in model 7 at $p < 0.01$). Thus, according to our results, we can support that the technological reputation that a firm gains by having developed a great number of new technological achievements in the past years appears to affect the breadth of its impact, but only when the breadth of impact is considered as the concentration of the technical classes of the influenced inventions (*Herfindahl Type Breadth)*. However, it is important to remind that we must be cautious in our interpretation of the results of *Recent Technological Activity*, since this variable is highly correlated with the control variable *Number of Patents*, causing an increase in the standards errors of those variables (Wooldridge, 2002).

Moreover, the findings in *Table 4.1* do not provide support for *Hypothesis 8a*, since the estimated coefficients of *Internal Focus* are insignificant in all the first three models. Recall that in *Hypothesis 8a* we had hypothesized that firms which build to a large extent upon their own technologies and neglect to absorb and assimilate the new technological developments introduced by other firms are in great danger of becoming technologically "isolated" and non-influential. However, our results do not indicate such a negative relationship between *Internal Focus* and technological impact. We suspect that the more a firm develops inventions that build on its own technologies, the deeper its knowledge on these technologies, and, therefore the higher the quality of these inventions. So, it might be the positive effect of this invention's quality on technological impact that counteracts the negative effect of technological isolation and thus makes technological impact to be independent of the *Internal Focus*. Similar findings have been reported in the study of Sapsalis et al. (2006) who also did not find any significant effect of self-backward citations on technological impact in the corporate sector.

Continuing with *Hypothesis 8b*, our findings reject this hypothesis, since all the estimated coefficients of *Internal Focus* are insignificant. Consequently, our results suggest that the breadth of technological impact, regardless of its type, is not associated with the degree to which a firm builds upon its prior R&D endeavors.

Regarding *Hypothesis 9a* and *9b*, generally the results reported in *Table 4.1* do provide full support for both of them. More specifically, all the estimated coefficients of R&D Investment are insignificant, suggesting that our response variables are probably unrelated to the amount of the investment in R&D. The only exception is the one-year lagged value of *R&D Investment* in model 3 which is positive and significant at $p < 0.1$.

However, this weak relationship between the one-year lagged value of *R&D Investment* and the *Technological Impact* becomes substantially strong when the sample is reduced only to those firms that receive a number of patent citations that is above the normal. In particular, instead of restricting the sample in the top 75th percentile, as we did in *Model 3*, we restricted the sample to those firm-year observations whose patent citations counts were larger than the average number of firm-year patent citations per year and per industry. As an example, the average number of firm-year patent citations in year *2003* in *Biotechnology* industry was 98. So, from *Biotechnology* industry in year 2003, we included only the firm-year observations whose patent citations counts were equal or higher than 98. We did the same for each year and for each industry. This model actually aims to examine the effect of our independent variables on the magnitude of impact only among the most important firms in terms of patent citations received. In this fixed-effects negative binomial model (observations: 221, Log likelihood = -731.29), the estimated coefficient of the lagged value of *R&D Investment* is 0.548 and it is significant at $p < 0.001$, suggesting that the more a firm invests in R&D, the greater the firm's impact on technological evolution, as measured by the aggregate number of citations received by its patents. The estimated coefficients of the rest of the explanatory variables are similar to *Model 3*[39].

[39] In *Appendix F*, we present the complete set of results of this model.

Finally, all the estimated coefficients of *Firm Size* are insignificant in all the first three models, and, thus, *Hypothesis 10a* is rejected. We believe that a reasonable explanation for this result is that the possible positive effects of firm size on technological impact such as reputation, prestige, and economies of scope and scale in research activities are eliminated by the negative effects of firm size such as bureaucracy, lack of entrepreneurial culture, and rigidity and inflexibility of knowledge base. Therefore, our findings suggest that the size of a firm does not have an impact on the degree to which it influences technological evolution. Furthermore, in contrast to our expectations, *Firm Size* does not seem to exert an influence on the breadth of technological impact, as we had hypothesized in *Hypothesis 10b*. Although, the estimated coefficients in model 8 may indicate some relationship, its significance (at $p < 0.1$) prohibits firm conclusions.

As far as the control variables are concerned, the coefficients of *Number of Patents* are, as expected, significantly positive in all models of technological impact (models 1-3) at $p < 0.0001$. In particular, *Number of Patents* exerts the strongest influence on our response variable since its coefficients are by far the highest of all. Furthermore, regarding the same control variable in the models of breadth of impact (models 4-8), our findings should come as no surprise, since the coefficients in all models are positive and significant (at $p < 0.\ 05$ in model 4, at $p < 0.0001$ in model 5, at $p < 0,001$ in model 7, and at $p < 0.1$ in models 6 and 8).

Moreover, the variable of *Number of Claims* appears to be an unimportant factor for the magnitude of impact, since the coefficients are insignificant in all the first three models. Trying to provide an explanation for these results, we suspect that the possible positive effect that the breadth of a patent with numerous claims has on its future citations is counteracted by the possible negative effect of the invention’s complexity. According to Leone and Reichstein (2012), higher number of claims denotes higher technological complexity, which may hamper the transfer of the knowledge that is embedded in the patent, by making more difficult for future patents to assimilate it. The same insignificance is present also in the models of breadth of impact as the estimated coefficients of *Number of Claims* in models 4-8 reveal.

With regard to effect of *Number of Inventors* on the magnitude of impact, as expected, the coefficients of *Number of Inventors* are significant and positive in models 1 ($p < 0.01$), 2 ($p < 0.05$), and 3 ($p < 0.05$), suggesting that the creative diversity of a patent, which is manifested through the number of its inventors, affects positively the citations that the patent receives (Nerkar 2003). Moreover, according to our results in models 4 and 5, the control variable of *Number of Inventors* seems to have no effect on the breath of technological impact. However, in models 6, 7, and 8, the findings provide some support for our expectations, since the estimated coefficients of *Number of Inventors* are positive and significant at $p < 0.01$ in models 7 and 8 and at $p < 0.10$ in model 6. Therefore, if the breadth of impact is measured as a Herfindahl Index, then the *Number of Inventors* appears to be an important factor.

Moreover, consistent with our expectations, all three models of technological impact suggest that both proxies of *Geographical Coverage* (*Patent Family Size* and *Designated States*) are important positive predictors of the magnitude of impact. Especially, the *Designated States* (i.e. the number of states or countries in which a patent family takes effect) appears to be strongly related to the dependent variable, since the estimated coefficients are positive and significant at $p < 0.001$ in model 1 and at $p < 0.01$ in models 2 and 3. The first proxy of *Geographical Coverage*, the *Patent Family Size*, appears to be less strongly related to the number of patent citations compared to *Designated States* (significant at $p < 0.10$ in models 1 and 2 and at $p < 0.05$ in model 3). Continuing with the models of breadth of impact, the coefficients of *Patent Family Size* and *Designated States* do not provide support for our prediction that the breadth of impact is positively related to *Geographical Coverage*. On the contrary, some coefficients suggest that the *Geographical Coverage* is negatively related to breadth of impact (*Patent Family Size* in model 6 at $p < 0.05$ and *Designated States* in models 4 and 6 at $p < 0.05$ as well).

Furthermore, in keeping with our expectations, the coefficients of *Number of Granted Patents* in all our models reveal a strong positive effect of *Number of Granted Patents* on patent citations ($p < 0.01$ in models 1 and 2, and $p < 0.001$ in model 3). However, this is not the case for the models of breadth of impact. In particular, the estimated coefficients of *Number of Granted Patents* are either insignificant (models 4-6) or positive but slightly significant at $p < 0.10$ (models 7 and 8).

With respect to the *Number of Citations Received* and counter to our expectations, the findings suggest that the *Herfindahl Type Breadth* is negatively related to the number of citations received. In particular, the estimated coefficients of *Number of Citations Received* in models 7 and 8 are significantly negative at $p < 0.01$ and at $p < 0.05$, respectively. The same results were reported by Banerjee and Cole (2010), who also used *Number of Citations Received* as a control variable and they found this variable to be negatively related to the Herfindahl type breadth.

Concerning the control variable *Profitability* in the first three models, the regression results suggest that *Profitability* exert a positive and significant effect on technological impact. More specifically, the estimated coefficients of the contemporaneous measures are positive and significant at $p < 0.01$, while the coefficients of the one-year lagged measures are positive and significant at $p < 0.05$ in all the first three models. Consequently, according to our findings, we can argue that the profitability of a firm is a positive and a significant predictor of its impact on the technological evolution. With regard to the models 4-8, generally, the findings suggest that firm profitability has no effect on the breadth of technological impact. The only exception is the one-year lagged value in model 5 (*Beyond-Firm Impact*), where the coefficient is positive and significant at $p < 0.05$. Thus, this result indicates that firm's profitability may have an impact on inventions that belong to different than the focal firm's technological classes.

Finally, as expected, the estimated coefficients of *Year* dummies in all the models suggest that that the age of a patent family affects positively both technological impact and breadth of impact, since the coefficients of the dummies that regard older years are always larger than the coefficients of the dummies that regard more recent years.

4.2 The Effect of the Magnitude and the Breadth of a Firm's Impact on its Financial Performance

At this point, a logical question emerges quite naturally. Do the firms who contribute more to the evolution of their technological environment enjoy better economic performance? It would be interesting to examine the extent to which the number of citations that a firm's patents receive is related to its economic performance.

Generally, the outcome of the innovative knowledge is widely believed to be a central determinant of a firm's competitive advantage and, thus, of its economic performance (Ceccagnoli 2009). According to He and Wang (2009), strong innovative knowledge strengthens a firm's ability to take advantage of opportunities in product markets and increases the likelihood of breakthrough inventions which, with an appropriate commercialization strategy, may increase a firm's rent generation potential. However, the competitive advantage that is gained by the innovative knowledge is transitory, because the new knowledge can spillover to rivals and quickly neutralize its positive effects on profitability. Firms, in order to cope with the spillover effect, use various strategies of rent appropriability, such as patenting, secrecy, first-mover advantages, and specialized complementary assets that can lead to higher price-cost margins and, eventually, to higher profitability (Ceccagnoli 2009).

More specifically with regard to patenting-firm performance link and especially for high influential patents, the extant literature suggests that the high impact of an invention may produce both a positive and a negative effect on firm performance. On the one hand, impactful patents can generate first-mover and network advantages that directly contribute to future earnings, while firms that possess influential patents are in a dominant position to persuade other firms to adopt their technological standards (Gu 2005). On the other hand, patents that receive a lot of citations may be exposed to imitation and consequently to the appropriation of their innovative knowledge by the rival firms, a fact which, in turn, may have a detrimental effect on the firm's financial performance (De Carolis 2003). In this case, impactful patented inventions, which probably require a great amount of resources - including highly R&D investments -, do not lead to substantial financial returns, making their contribution to the firm's financial performance insignificant or even negative. A prominent example of knowledge imitability is Xerox, whose Palo Alto Research Center (PARC) developed many impactful technologies that were used in personal computing, graphical displays, and computer networking but failed to capitalize its inventions.

The relevant literature on patent citations - firm performance link provides mixed empirical evidence regarding the impact of patent citations on firm's financial performance, even if one can claim that the studies that present a positive relationship are more than the studies that present a negative relationship. For example, He and Wang

(2009) presented a positive and significant effect of the total patent citations that a firm receives on its economic performance. In addition, Markman et al. (2004) found that highly cited patents are positively related to superior performance, while Gu (2005) showed that an increase in patent citation impact of 1% is associated with an increase in earnings of the next year by 0.111%. More evidence for the statistically and economically significant relation of patent citations and firm performance can be found in the study of Hall et al. (2005), who found that on average, for every one additional citation that a firm's patent receives, the firm's market value increases by 3%.

On the contrary, Chen and Chang (2010) showed that patent citations have an inverse U-shaped influence upon corporate market value in the pharmaceutical industry, and, in addition, that there exists an optimal value beyond which patent citations are negatively associated with corporate market value, mainly because of the significant spillover effect of innovative knowledge. On the basis of this evidence, the authors suggested that "if a pharmaceutical company does have more patent citations beyond the optimum level, the spillover effect of its R&D knowledge is so significant that it becomes difficult to capture enough appropriable rights from its patents to secure a profit". Similarly, Branch and Chichirau (2010) reported a negative association between patent citations and profitability and concluded, among others, that R&D is oriented toward the long-term profitability as well as being expensive and risky with a volatile outcome.

It is our expectation that the firms, whose patents receive a substantial number of citations, will achieve better financial performance as compared to the firms whose patents receive relatively few citations. The key point in this relation is the ***profit margin*** that it can be increased when the firm considers that the higher prices of a certain product (or service) will not cause the reduction of the product sales, based on the belief that the customers are willing to pay a higher price than to change product. Thus, the substantial number of citations that the patents of a firm receive can be interpreted as an indication that there exist numerous other technologies that build upon and are locked in the focal firm's technologies and probably will continue to follow these technologies even if their cost increase. However, we have to emphasize that the above argument is build on the assumption that the technologies that are locked in others technologies in terms of patented inventions and patent citations are also locked in those technologies in terms of products (or services). In other words, the assumption is that the dependence of a technology on another technology is present not only in the patent system but also in the market.

However, we anticipate that in some cases the substantial number of citations that the patents of a firm receive, may turn out to be a negative factor "in disguise". Imagine, for example, a firm that has developed a patented invention which has received a great amount of patent citations. The acceptance of this invention may cause the managers of the firm to decide to commercialize it, in anticipation of high financial returns. So, the firm launches a new business plan, based on this invention, and increases its expenditure regarding the development and the marketing of the product (or the service) that contains the technology of this specific invention. However, rivals that have developed inventions that are partly rely on this invention (as manifested by patent citations) will have probably assimilated this technical knowledge and they may have proceed to the development of superior and commercially more successful products as compared to the product of the first firm. So, in these cases, the investment will not lead to the anticipated profit and will affect negatively the financial performance. Nevertheless, we expect that the negative aspects of the patent citations with regard to financial performance will not prevail against or neutralize completely its positive aspects. Therefore, all the above considerations lead us to the following hypothesis:

Hypothesis 11a: *Ceteris paribus*, the greater the impact of a firm on technological evolution, as measured by the aggregate number of citations received by its patents, the higher its financial performance.

As far as the relation between the breadth of impact and the financial performance is concerned, the extant literature, to our knowledge, does not offer any empirical evidence on this issue. Nevertheless, we expect that the firms that realize that their patented inventions have a broad influence on subsequent patents – by monitoring the technological origins of the patents that cite their patents - will make an effort to commercially exploit this particular characteristic of their technology. Probably, the managers of the firms will try to create a new competitive advantage on the basis of the technologies that present a high breadth of impact. For example, a firm with substantial breadth of impact could launch a joint venture with a partner that operates in a technological domain in which the focal firm is not involved but it has an important influence on. In particular, the focal firm may form a joint venture with a firm from a different technological domain that has build part of its technology on the focal firm's technologies. This joint venture could aim at the production of an innovative product or a series of innovative products that are new to the market that, in turn, may provide a competitive advantage for the joint venture and lead to above-normal financial returns. Thus, based on the above argumentation, we propose our next hypothesis:

Hypothesis 11b: *Ceteris paribus*, the wider the breadth of a firm's technological impact, the higher its financial performance.

However, it is important to note that this particular effect of breadth of impact on financial performance that we just mentioned, it might be difficult to appear in the short-run, since the formation of a joint venture and the development of innovative products that are based on various technologies from different technological areas are long-term processes. Therefore, because our empirical analysis examines the causal effect of breadth of impact on financial performance up two years lag, it is likely that this effect will not emerge.

We used *Operating Profit Margin* (*Net Income to Sales ratio* or *Return on Sales*), in order to compute our dependent variable (Robinson and McDougall 1998). *Operating Profit Margin* is expressed as a percentage and, in effect, measures how much out of sales a company actually keeps in earnings. We used *Operating Profit* and *Net Sales* from *Scoreboard* in order to compute *Operating Profit Margin*.

In our regression models, apart from the main independent variables that we want to examine their impact on financial performance, that is to say the magnitude of impact and the three types of breadth of impact (*Beyond-Industry Impact*, *Beyond-Firm Impact*, and *Herfindahl Type Breadth*), we used some other firm characteristics as control variables. More specifically, we use the contemporaneous value and the one-year lagged values of *Net Sales*, *R&D Investment*, *Number of Employees*, *Capital Expenditures (Capex)*[40], since all of these variables may exert an influence on *Operating Profit Margin.* The source of the aforementioned variables is the *Scoreboard*.

But, since there exist plenty of significant factors, that are not included in the model and affect firm financial performance –e.g. marketing competencies (De Carolis 2003), competitive manufacturing, after sales support (Rothaermel and Hill 2005),

[40] Capital expenditure (Capex) is expenditure used by a company to acquire or upgrade physical assets such as equipment, property, industrial buildings. We acquired data for Capex from the *EU Industrial R&D Investment Scoreboard*.

organizational structure, human resources (Koellinger 2008)–, we proceed to the inclusion of the lagged dependent variable, in order to capture the possible effects of the unobserved variables on our dependent variable. However, the inclusion of the lagged dependent variable causes the unobserved panel-level effects to be correlated with the lagged dependent variables, making standard estimators inconsistent (Benner and Ranganathan 2012). To address this issue, we used an improved version of the Arellano and Bond's (1991) generalized method-of-moments (GMM) estimator, the Blundell–Bond linear dynamic panel data estimator (the *xtdpdsys* command in Stata), which is designed for datasets with many panels and few periods (Arellano and Bover, 1995, Blundell and Bond, 1998).

Since, the hypothesized effect of magnitude of impact or of breadth of impact on financial performance takes place always with a substantial time-lag, we used the two-year lagged value of our independent variables apart from the one-year lagged value. Time invariant variables, such as industry or country, are omitted from the Blundell-Bond estimation because the model uses the method of first differences. Finally, we have to mention that we included year dummies and we controlled for the number of patents in order to avoid any potential bias between citations and financial performance due to the variability of the number of patents.

In *Table 5.2*, we present the results of our regression models. Models 9 and 10 test the effect of one-year lagged value of magnitude of impact and the two-year lagged, respectively, on financial performance (*Hypothesis 11a*). In models 11, 13, and 15, we examine the impact of one-year lagged value of *Beyond-Industry Impact*, *Beyond-Firm Impact*, and *Herfindahl Type Breadth*, respectively, on *Operating Profit Margin*, while in models 12, 14, and 16 we test the effect of the two-year lagged value of these variables (*Hypothesis 11b*).

Moreover, because *xtdpdsys* command assumes that there is no autocorrelation in the idiosyncratic errors, we used a specification test (Cameron and Trivedi 2009) to ensure this. Indeed, the *estat abond* command in *STATA* showed that the idiosyncratic errors are serially uncorrelated (the null hypothesis that are serially correlated is rejected at order 1 at $p < 0.10$ in all models). An additional specification test is a test of overidentifying restrictions. This test builds a null hypothesis that the overidentifying restrictions are valid. A possible rejection of the null hypothesis implies that we need to reconsider our model or our instruments. The *estat sargan* command is *STATA* implements this test (Cameron and Trivedi 2009). In our case, the *estat sargan* tests did not cast any doubt on the validity of the instruments used, since in none of the models presented strong evidence against the null hypothesis that the overidentifying restrictions are valid ($p > 0.096$ in all models). As an assessment of the goodness-of-fit, *STATA* provides the value of the sum of squared differenced residuals (RSS). The sum of squared differenced residuals measures the discrepancy between the data and the estimation model. Low values denote a good fit of the model to the data. *Model 10* appears to have the best fit among *Models 9-16* since its RSS value is the lowest of all. (*Model 9*: 154.90, *Model 10*: 128.21, *Model 11*: 174.43, *Model 12*: 145.78, *Model 13*: 174.59, *Model 14*: 144.12, *Model 15*: 171.63, *Model 16*: 143.41)

The findings in *Table 5.2* suggest that the firm's impact on technological evolution is unrelated to its future financial performance, since the estimated coefficients of *Lagged Magnitude of Impact* and *Two-year Lagged Magnitude of Impact* in models 9 and 10 are insignificant. It is worth noting that the estimated coefficient of *Lagged Magnitude of Impact* in model 9 is negative but significant only at $p < 0.10$, a level of significance that is preventing us from drawing any firm conclusion. Consequently, we believe that the potential positive effects of being technological influential are counteracted by the negative effects of imitation and spillover and, thus, the extent to which a firm influences the technological evolution appears to be unrelated to its financial performance. The same results were obtained in the case of breadth of impact. More specifically, the estimated coefficients of models 11-16 are all insignificant, leading us to argue that the firm's breadth of impact is not associated with its future financial returns.

Furthermore, a series of robustness checks were conducted. First, we used consecutively *Operating Profit* and *Sales* instead of *Operating Profit Margin* as our dependent variable, examining if different measures of financial performance can create different relations with our dependent variables. In both of these cases, no relation between magnitude or breadth of impact and financial performance was found. Second, we tested the effect of the three-year lagged value of our explanatory variables on *Operating Profit Margin*. Again, the estimated coefficients were insignificant. Finally, we examined if there is any inverted-U shaped type of relation between our explanatory and response variables, but results did not suggest any such type.

Table 5.2 The Blundell-Bond Linear Dynamic Panel Data Model of the Effect of a Firm's Technological Impact or Breadth of Impact on its Financial Performance

Variables	Model 9 Lagged Magnitude of Impact	Model 10 Two-year Lagged Magnitude of Impact	Model 11 Lagged Beyond-Industry Impact	Model 12 Two-year Lagged Beyond-Industry Impact	Model 13 Lagged Beyond-Firm Impact	Model 14 Two-year Lagged Beyond-Firm Impact	Model 15 Lagged Herfindahl Type Breadth	Model 16 Two-year Lagged Herfindahl Type Breadth
Lagged Firm Performance	0.015 (0.109)	-0.001 (0.210)	-0.094 (0.126)	-0.068 (0.110)	-0.094 (0.128)	-0.104 (0.104)	-0.112 (0.128)	-0.092 (0.111)
Lagged Technological Impact	-0.278+ (0.167)	-0.226 (0.207)						
Two-year Lagged Technological Impact		0.027 (0.222)						
Lagged Beyond-Industry Impact			-0.001 (0.025)	0.002 (0.025)				
Two-year Lagged Beyond-Industry Impact				-0.038 (0.026)				

Variables	Model 9 Lagged Magnitude of Impact	Model 10 Two-year Lagged Magnitude of Impact	Model 11 Lagged Beyond-Industry Impact	Model 12 Two-year Lagged Beyond-Industry Impact	Model 13 Lagged Beyond-Firm Impact	Model 14 Two-year Lagged Beyond-Firm Impact	Model 15 Lagged Herfindahl Type Breadth	Model 16 Two-year Lagged Herfindahl Type Breadth
Lagged Beyond-Firm Impact					-0.037 (0.043)	-0.012 (0.052)		
Two-year Lagged Beyond-Firm Impact						0.041 (0.066)		
Lagged Herfindahl Type Breadth							-0.288 (0.211)	-0.135 (0.266)
Two-year Lagged Herfindahl Type Breadth								0.158 (0.345)
Lagged Number of Patents	0.265 (0.219)	0.109 (0.249)	-0.178 (0.224)	-0.093 (0.155)	-0.169 (0.221)	-0.135 (0.160)	-0.105 (0.192)	-0.100 (0.161)
Two-year Lagged Number of Patents		0.219 (0.337)		0.085 (0.230)		0.067 (0.249)		0.013 (0.213)
Sales	2.335* (0.947)	2.030+ -1.234	1.431+ (0.808)	1.664+ (0.862)	1.408+ (0.811)	1.728* (0.868)	1.446+ (0.802)	1.658* (0.795)
Lagged Sales	0.366	-0.084	0.335	-0.184	0.310	-0.066	0.367	-0.182

Variables	**Model 9** Lagged Magnitude of Impact	**Model 10** Two-year Lagged Magnitude of Impact	**Model 11** Lagged Beyond-Industry Impact	**Model 12** Two-year Lagged Beyond-Industry Impact	**Model 13** Lagged Beyond-Firm Impact	**Model 14** Two-year Lagged Beyond-Firm Impact	**Model 15** Lagged Herfindahl Type Breadth	**Model 16** Two-year Lagged Herfindahl Type Breadth
	(0.873)	-1.043	(0.810)	(0.783)	(0.841)	(0.800)	(0.774)	(0.784)
R&D Investment	-0.802* (0.316)	-0.708* (0.340)	-0.529* (0.265)	-0.507+ (0.302)	-0.529* (0.269)	-0.500 (0.313)	-0.567* (0.266)	-0.501 (0.309)
Lagged R&D Investment	0.432 (0.377)	0.363 (0.483)	0.062 (0.285)	0.088 (0.280)	0.067 (0.279)	0.059 (0.287)	0.099 (0.267)	0.150 (0.266)
Firm Size	-0.327 (0.662)	-0.202 -1.329	0.198 (0.665)	0.174 (0.646)	0.286 (0.656)	0.224 (0.633)	0.256 (0.678)	0.149 (0.658)
Lagged Firm Size	-1.556* (0.774)	-1.021 -1.018	-1.277+ (0.772)	-0.937 (0.694)	-1.296 (0.793)	-1.088 (0.727)	-1.421+ (0.822)	-0.993 (0.713)
Capital Expenditure	0.053 (0.210)	-0.071 (0.278)	-0.234 (0.214)	-0.272 (0.266)	-0.257 (0.217)	-0.285 (0.267)	-0.187 (0.206)	-0.222 (0.275)
Lagged Capital Expenditure	0.039 (0.222)	-0.007 (0.320)	0.082 (0.206)	0.038 (0.232)	0.071 (0.209)	0.034 (0.233)	0.094 (0.204)	0.027 (0.248)
Year Dummies Included								

Variables	**Model 9** Lagged Magnitude of Impact	**Model 10** Two-year Lagged Magnitude of Impact	**Model 11** Lagged Beyond-Industry Impact	**Model 12** Two-year Lagged Beyond-Industry Impact	**Model 13** Lagged Beyond-Firm Impact	**Model 14** Two-year Lagged Beyond-Firm Impact	**Model 15** Lagged Herfindahl Type Breadth	**Model 16** Two-year Lagged Herfindahl Type Breadth
_cons	1.575*** (0.442)	1.781* (0.731)	2.320**** (0.450)	2.290**** (0.350)	2.296**** (0.445)	2.364**** (0.347)	2.492**** (0.485)	2.328**** (0.519)
Sum of Squared Differenced Residuals	154.90	128.21	174.43	145.78	174.59	144.12	171.63	143.41
N	529	440	554	464	554	464	554	464
Sargan test	chi2(19) = 23.34541 Prob > chi2 = 0.2225	chi2(18) =25.30557 Prob > chi2 = 0.1167	chi2(19 = 26.01962 Prob > chi2 = 0.1296	chi2(18) = 23.97658 Prob > chi2 = 0.1558	chi2(19) = 26.15826 Prob > chi2 = 0.1258	chi2(18) = 24.64173 Prob > chi2 = 0.1351	chi2(19) = 26.59444 Prob > chi2 = 0.1145	chi2(18) = 26.16131 Prob > chi2 = 0.0961
Estat abond - order 1	0.0157	0.0583	0.0555	0.0786	0.0572	0.0807	0.0718	0.0851
Estat abond - order 2	0.1344	0.1533	0.1021	0.1592	0.1033	0.1428	0.1067	0.1622
Estat abond - order 3	0.2736	0.0934	0.4682	0.4140	0.4558	0.4110	0.4839	0.4117

Variables	Model 9 Lagged Magnitude of Impact	Model 10 Two-year Lagged Magnitude of Impact	Model 11 Lagged Beyond-Industry Impact	Model 12 Two-year Lagged Beyond-Industry Impact	Model 13 Lagged Beyond-Firm Impact	Model 14 Two-year Lagged Beyond-Firm Impact	Model 15 Lagged Herfindahl Type Breadth	Model 16 Two-year Lagged Herfindahl Type Breadth
Standard errors in parentheses **+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001**								

CHAPTER 5

DISCUSSON AND CONCLUSIONS

5 DISCUSSION AND CONCLUSIONS

5.1 Introduction

This study tackled an important issue in organization theory regarding the general issue of the relationship between organizations and their environment. We endorsed a broadly co-evolutionary framework where organizations are constrained by their environment but are also endowed with an ability to change it. We focused specifically on the organizations' ability to change their technological environment, and we use patent citations as a proxy variable that reflects the impact of an organization on technological evolution.

The main objective of this study was to explore the ***firm-level factors*** that affect the extent to which a firm shapes technological change and exerts a broad influence on it. We proposed and we tested ten hypotheses with respect to those factors. Moreover, we attempted to explore the effect of the magnitude of impact as well as the breadth of impact on financial performance.

Methodologically, this study used longitudinal (from 2003 to 2009), multi-industry (139 firms from biotechnology, pharmaceuticals, chemicals), multi-national (19 countries of origin) secondary data from two sources, the *EU Industrial R&D Investment Scoreboard* and the *Derwent Innovation Index Database*. The former source provided us with "purely" firm-level characteristics (e.g. R&D expenditure, sales, number of employees etc.) while the latter provided us with patent information on the basis of which we attempted to create proxies for a series of firm characteristics (e.g. technological collaborations, technological diversity, recent technological activity etc.)

5.2 Main Findings

With respect to the firm-level factors that affect the ***magnitude of impact***, our empirical findings provided support for some –but not all- of our hypotheses. Importantly, and in line with our predictions, we found that the more recent the technologies a firm incorporates into its inventions, the greater its impact on technological evolution. It is our belief that this effect is probably due to the fact that in recent technologies, an inventor can discover a higher number of useful combinations of the elements of these technologies, compared to mature technologies. These useful combinations can, in turn, lead to influential inventions (Heeley and Jacobson 2008). On the contrary, the technological inventions that rely on the recombination of mature technological components may not be able to provide successful technological solutions to the *current* technological challenges which are probable fundamentally different from those mature technologies.

Surprisingly, and contrary to our expectations, our empirical results showed that that the firms that experiment with unexplored technological areas to a large extent and, presumably, develop inventions that lack technological antecedents, are less able to influence technological evolution compared to firms that rely heavily on existing knowledge and develop incremental inventions. It seems that, in an aggregate sense, whenever the inventor of a future invention confronts the dilemma of choosing between a radical technology, from which an enormous technological success may emerge but, as untested, may also result in a complete failure and an incremental technology with which the inventor is familiar and in which the inventor is able to find and to assess those components that can be useful to him/her, it prefers the security and the estimable success of the latter technology.

It is also worth mentioning that, as the empirical results showed, the more narrow the technological roots upon which the inventions are built, the greater their magnitude of impact. Relying on just a few technological domains reinforces the depth of specialization in these domains and the inventor may be capable of developing profound technological solutions. On the contrary, when an invention relies on wide and numerous technological domains, it is likely that although this invention is radical in terms of the novelty of the combinations of the knowledge components, its generality of its technological roots probably weakens its expertise and the solution it provides are probably of less quality in comparison to the specialized inventions (i.e. inventions that rely on a narrow range of technological domains). Moreover, on the part of the inventors of the future patents, it must be again the inconvenience they feel whenever they face the unfamiliar, the unknown and the uncertainty that characterized usually the radical inventions, as was the case with the unexplored technological areas, Therefore, firms that create inventions of a higher degree of specialization appear to have an increased ability to shape technological evolution.

Apart from the effects that were found significant, some interesting conclusions can be derived also from the insignificant effects. For example, with respect to the independence of technological impact from the extent to which a firm's R&D builds upon its prior technological developments, we conjecture that the potential positive effects of deepening the knowledge gained in previous inventions is counteracted by the potential negative effects of the technological isolation that, almost inevitably, accompanies internal technological focus.

Additionally, a syllogism that can interpret the absence of any relationship between technological impact and firm size, is the following: The net effect of firm size is insignificant because as the size of a firm increases, on the one hand, the social power of the firm strengthens (higher reputation and prestige, enhanced ability to put pressure on governments for favorable regulations) and the benefits of the economies of scale and scope are becoming greater, but, on the other hand, the bureaucracy, the lack of entrepreneurial culture, and the organizational rigidity increases as well. Therefore, these antithetic sets of forces neutralize any possible individual effect that each force separately may exert on the magnitude of impact.

As far as the firm-level factors affecting the ***breadth of technological impact*** are concerned, our empirical results revealed some notable findings. First, with regard to the firm's impact on industries others than its own (***Beyond-Industry Impact***), we found evidence that scientific knowledge stemming from patents that have been applied in cooperation with universities exerts a positive influence on *Beyond-Industry Impact*. We attribute this finding to the theoretical knowledge, which characterizes scientific inventions, that is by nature more general than the practical knowledge of technological inventions. Moreover, we found that the breadth of impact that pertains to a firm's technological impact outside of the boundaries of its industry is affected negatively by the degree to which the firm develops inventions that are categorized in many technological domains. It is our belief that inventions that "inject" knowledge from a wide set of technological domains probably provide technological solutions that are superficial, in contrast to the inventions that rely on a narrow set of technological domains which presumably deepen the knowledge within these domains and achieve to create specialized technical knowledge so profound that may encompasses pieces of knowledge with a broad applicability. As an example, when the R&D unit of a chemical firm develops an invention that deepens the existing knowledge in specific chemical technological fields, a new piece of knowledge may be created that will regard the science of chemistry in general (and not just the chemical products) and it might be useful in industries that are partly based on chemistry such as biotechnology or pharmaceuticals.

Second, with respect to the firm's impact on technological fields others than the ones that the firm's operates in (***Beyond-Firm Impact***), we can say that it is very difficult to detect any firm level characteristic that can have a causal effect on this type of breadth. Almost all the firm level factors examined do not appear to be casually related to *Beyond-Firm Impact.* The only exception appears to be the scientific basis of the firm's inventions, as in the case of *Beyond-Industry Impact*. The reliance on scientific knowledge presumably results in the creation of new pieces of technical knowledge with a strong scientific nature. These pieces of "scientific technical knowledge" seem to be useful in different technological domains than those in which they have been developed, rather because of their generality in their applicability.

Third, concerning the third type of breadth, the ***Herfindahl Type Breadth***, we found that firms that are based to a large extent on scientific knowledge are exerting a more broad technological influence. This effect is in agreement with the other two types of breadth of impact and, therefore, we can conclude that the more a firm relies on the scientific sphere to develop new technologies, the higher its breadth of impact regardless of its type. However, technological diversity, which was found to affect negatively *Beyond-Industry Impact*, in the case of *Herfindahl Type Breadth* was found to be positively related. Putting these two types of breadth together, it seems that as the technological bases of an invention become more diverse, the affected technological domains become more diverse also (i.e. less concentrated) but their probability to belong to different industry from the focal invention’s industry decreases. Moreover, it must be the high technological reputation that results from the intense prior technological activity of a firm that makes their inventions more attractive to a broader range of technological areas.

Concerning the independence of magnitude and breadth of impact from financial performance, we believe that main reason for the absence of any relationship is the inability of the impactful firms to “transfer” their degree of influence from the sphere of technological inventions to the sphere of market and products. Maybe it is the problematic development of the product which although it incorporates influential technological inventions, it cannot cope successfully with the market needs or with the competition in specific market segments. After all, in the market arena, the technological superiority cannot stand by itself without be combined with other significant parameters such as the marketing strategy, the product design, the distribution channels, and so on. Or, it could be the spillover effect that can give the opportunity to rivals to assimilate the new technology and to develop an alternative, improved version of this technology and in combination with a significant R&D-manufacturing-marketing time lag in the focal firm can lead to the elimination of the potential financial benefits of being technologically influential.

Moreover, it seems that firms do not exploit their ability to develop technologies with a substantial breadth of impact. Although they can monitor the transfer of their novel knowledge components to other technological domains, and presumably, to other market segments, they probably do take any action to capitalize their advantage by forming joint ventures, for example, with firms from those segments, aiming at the creation of a differentiated, innovative product that can lead to high profit margins.

5.3 Limitations of the Study

Our study is limited in many respects. First, patents provide only a partial measure of the inventive activity of a firm. In particular, although we may know the exact number of patents that were applied by a firm, we cannot know the exact proportion of the total inventions of the firm that these patents represent (Hoetker and Agarwal 2007). The reason is that only a part of the firm's inventions are successful enough to have resulted in patents (Rosenkopf and Nerkar 2001). Second, for the same reason, patent citations can be considered as an incomplete measure of knowledge transfer. For example, in the cases of inventions that are not resulted in a patentable technology, all the prior knowledge that the inventions have build upon cannot be tractable via patent citations. Third, whenever an analysis focuses on certain industries, results may not be generalizable to other industries that have very different conditions (Cattani 2005, Hoetker and Agarwal 2007). Our results are more likely to apply in high-tech industries with a strong regulatory context. Fourth, since our sample consists of large firms with high R&D expenditure, all the statistical inferences regard mainly the large firms of our focal industries and not just any firm of these industries.

5.4 Managerial Implications of the Study

The findings of this study may have important implications not only for the understanding, by the scholars, of the factors that affect the extent to which a firm shapes technological evolution, but also may have noteworthy managerial implications. The managers of the firms with high R&D expenditure that operate in high-tech industries need to be cognizant about the determinants of the firm's ability to influence technological change, since this ability may prove to be very beneficial for them. In particular, the systematic adoption by other firms of a firm's technologies will probably increase the likelihood of firm's survival and growth and may result in better financial performance in the medium or in the long-term (regarding the short-term effect, our model showed that it is insignificant).

Moreover, with regard to the implications that concern the breadth of technological impact, scholars can, aided by this study, detect the firms which are responsible for the connectivity of different technological fields and attribute this ability to certain firm characteristics. In addition, managers with the knowledge of this study can focus on these firm-level competencies that increase the breadth of technological impact, in anticipation of the tangible benefits that may result from a potential innovative cooperation (e.g. joint venture) with firms from different technological domains, which, in turn, may lead to product differentiation as well as to product diversification and to economies of scope, and, consequently, to firm efficiency and effectiveness.

Furthermore, the appropriability regime in general, and the patent system in specific, cannot provide any benefits to those technologies that have been partly used, together with other technologies, in the development of new technologies. Because, although these new technologies have employed some pieces of knowledge from previous inventions, they are so different from them, that they cannot be accused of violating properties rights, especially if the technological domain of the new technology belongs to different industry and focuses on different market segments than the technologies that is based on. Therefore, in such cases where an impactful firm sees its technological inventions to exert a great influence on future technologies but it cannot claim for any property right or for any financial gain, the managers of the influential firm could approach the firms of the "affected technologies" and to start a conversation for a collaborative perspective on the basis of the strong affinity of their technological interests that may lead to technological synergies and to successful innovative products. So, the managers by monitoring the flow of their forward patent citations, they could plan various strategic scenarios aiming at some sort of collaboration with the firms that use their technologies.

5.5 Recommendations for Further Study

As we aforementioned, one important limitation of our study is that our results cannot be fully generalizable to other industries. At the same time, this limitation offers scope for future research. In particular, future efforts comparing the technological evolution of different industries will verify whether the same relations between our response and our explanatory variables can be detected. In addition, future researchers could enrich our econometric models by examining the effect of new explanatory variables on our response variables. For example, it would be interesting to see if the marketing activities of a firm are related to the extent to which its technology is been accepted by its technological community, or if the mergers and acquisitions can have a direct effect on the degree to which a firm shapes technological evolution.

Moreover, another possible future avenue of this research is to understand more comprehensively how the firm’s ability to shape technological evolution is related to its economic performance. For example, the econometric model of firm performance could be extended to include more factors that exert influence on firm performance (such as marketing competencies, organizational structure and human resources). This will produce a better estimate of the effect of technological impact on financial performance. Furthermore, it would be interesting to test this relation not only in the short term (we tested the one-year and the two-year lagged effect), but also in the long-term. We suspect that the positive effect of technological impact on financial performance in the long-term could be empirically proven.

Additionally, it would be important to examine how much of the magnitude of impact is attributable to technological superiority and how much to the social power. This could be partly achieved by testing the effect of variables that contribute exclusively to social power such as the lobbying expenses (can influence regulations) or the marketing expenses (can influence cognition, reputation, trends etc.) on the magnitude and the breadth of technological impact. Moreover, all these above-mentioned variables could be used in the creation of an index of the "magnitude of institutional impact", analogous to the magnitude of technological impact, in an effort to enhance the research area that aims to examine all the possible aspects of the ability of an organization to influence the evolution of its environment in general and the factors that can affect this ability.

Another interesting future research avenue can emerged from the matching of the patented inventions with the innovative products that have resulted from these inventions. For example, one could examine if the magnitude or the breadth of impact of these patented inventions are related to the sales, the market shares and the profits of these specific innovative products. Moreover, it would be also important to analyze and compare the market performance of an innovative product that has been derived from certain patented inventions and other innovative product that have been derived from patented inventions that to a large extent rely on the focal patent invention. This comparison could enlighten us concerning the issue that regards the spillover effect and its impact on the realm of products and markets.

Finally, is seems important to examine if there is any relationship between the organizations of a joint venture or R&D consortia and their connectivity in terms of patent citations. In particular, scholars could analyze if the organizations of the influential patents are collaborating technologically in any way with the organizations of the "affected" patents.

BIBLIOGRAPHY

Agarwal, R., M. Ganco, R.H. Ziedonis. 2009. Reputations for toughness in patent enforcement: implications for knowledge spillovers via inventor mobility. *Strategic Management Journal* **30**(13) 1349-1374.

Ahuja, G. 2000. Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly* **45**(3) 425-455.

Ahuja, G., C. Morris Lampert. 2001. Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal* **22**(6β€ 7) 521-543.

Alcacer, J., M. Gittelman, B. Sampat. 2009. Applicant and examiner citations in US patents: An overview and analysis. *Research Policy* **38**(2) 415-427.

Alencar, M.S.M., A.L. Porter, A.M.S. Antunes. 2007. Nanopatenting patterns in relation to product life cycle. *Technological Forecasting and Social Change* **74**(9) 1661-1680.

Allison, P. D. 2009. *Fixed effects regression models*. SAGE publications.

Allison, P. D., R. P. Waterman. 2002. Fixed-effects negative binomial regression models. *Sociological Methodology* **32**(1) 247-265.

Anand, J., R. Oriani, R.S. Vassolo. 2010. Alliance activity as a dynamic capability in the face of a discontinuous technological change. *Organization Science* **21**(6) 1213-1232.

Arellano, M., S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies* **58**(2) 277-297.

Arellano, M., O. Bover. 1995. Another look at the instrumental variable estimation of error-components models. *Journal of econometrics* **68**(1) 29-51.

Argyres, N.S., B.S. Silverman. 2004. R&D, organization structure, and the development of corporate technological knowledge. *Strategic Management Journal* 25 929-958.

Astley, W.G. 1985. The two ecologies: Population and community perspectives on organizational evolution. *Administrative Science Quarterly* 224-241.

Bailey, D.E., S.R. Barley. 2011. Teaching-learning ecologies: Mapping the environment to structure through action. *Organization science* **22**(1) 262-285.

BIBLIOGRAPHY

Banerjee, P. M., B. M. Cole. 2010. Breadth-of-impact frontier: How firm-level decisions and selection environment dynamics generate boundary-spanning inventions. Technovation **30**(7) 411-419.

Barley, S.R., P.S. Tolbert. 1997. Institutionalization and Structuration: Studying the Links between Action and Institution. *Organization Studies* **18**(1) 93-117.

Battilana, J., B. Leca, E. Boxenbaum. 2009. How Actors Change Institutions: Towards a Theory of Institutional Entrepreneurship. *The Academy of Management Annals* **3**(1) 65-107.

Baum, J.A.C., J.V. Singh. 1994a. Organizational hierarchies and evolutionary processes: Some reflections on a theory of organizational evolution. J.A.C. Baum, J.V. Singh, eds. *Evolutionary dynamics of organizations*. Oxford University Press, 3-20.

Baum, J.A.C., J.V. Singh. 1994b. Organization-environment coevolution. J.A.C. Baum, J.V. Singh, eds. *Evolutionary dynamics of organizations*. Oxford University Press, 379-402.

Bekkers, R., G. Duysters, B. Verspagen. 2002. Intellectual property rights, strategic technology agreements and market structure: The case of GSM. *Research Policy* **31**(7) 1141-1161.

Benner, M.J., R. Ranganathan. 2012. Offsetting illegitimacy? How pressures from securities analysts influence incumbents in the face of new technologies. *Academy of Management Journal* **55**(1) 213-233.

Benner, M.J., M. Tushman. 2002. Process management and technological innovation: A longitudinal study of the photography and paint industries. *Administrative Science Quarterly* **47**(4) 676-707.

Bessen, J. 2008. The value of US patents by owner and patent characteristics. *Research Policy* **37**(5) 932-945.

Blundell, R., S. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics* **87**(1) 115-143.

Branch, B., C. Chichirau. 2010. Mispricing vs risk premia in R&D-intensive firms. *International Review of Financial Analysis* **19**(5) 358-367.

Cameron, A.C., P.K. Trivedi. 2009. *Microeconometrics using stata*. Stata Press College Station, TX.

Cantwell, J., J.H. Dunning, S.M. Lundan. 2009. An evolutionary approach to understanding international business activity: The co-evolution of MNEs and the institutional environment. *Journal of International Business Studies* **41**(4) 567-586.

BIBLIOGRAPHY

Carney, M., E. Gedajlovic. 2002. The co-evolution of institutional environments and organizational strategies: The rise of family business groups in the ASEAN region. *Organization Studies* **23**(1) 1-29.

Cattani, G. 2005. Preadaptation, firm heterogeneity, and technological performance: a study on the evolution of fiber optics, 1970-1995. *Organization science* **16**(6) 563-580.

Caves, R.E., M.E. Porter, A.M. Spence. 1980. *Competition in the open economy: A model applied to Canada*. Harvard University Press.

Ceccagnoli, M. 2009. Appropriability, preemption, and firm performance. *Strategic Management Journal* **30**(1) 81-98.

Chen, Y.-S., K.-C. Chang. 2010. Exploring the nonlinear effects of patent citations, patent share and relative patent position on market value in the US pharmaceutical industry. *Technology Analysis & Strategic Management* **22**(2) 153-169.

Child, J., Y. Lu, T. Tsai. 2007. Institutional entrepreneurship in building an environmental protection system for the People's Republic of China. *Organization Studies* **28**(7) 1013-1034.

Child, J., T. Tsai. 2005. The Dynamic Between Firms' Environmental Strategies and Institutional Constraints in Emerging Economies: Evidence from China and Taiwan. *Journal of Management Studies* **42**(1) 95-125.

Dahlin, K.B., D.M. Behrens. 2005. When is an invention really radical? Defining and measuring technological radicalness. *Research Policy* **34**(5) 717-737.

De Carolis, D.M. 2003. Competencies and imitability in the pharmaceutical industry: An analysis of their relationship with firm performance. *Journal of Management* **29**(1) 27-50.

Dieleman, M., W.M. Sachs. 2008. Coevolution of institutions and corporations in emerging economies: How the Salim group morphed into an institution of Suharto's crony regime. *Journal of Management Studies* **45**(7) 1274-1300.

DiMaggio, P.J. 1988. Interest and agency in institutional theory. L. Zucker, ed. *Institutional patterns and organizations: Culture and Environment*. MA: Ballinger, 3–22.

Dushnitsky, G., M.J. Lenox. 2005. When do incumbents learn from entrepreneurial ventures?: Corporate venture capital and investing firm innovation rates. *Research Policy* **34**(5) 615-639.

Fleming, L. 2001. Recombinant uncertainty in technological search. *Management Science* **47**(1) 117-132.

BIBLIOGRAPHY

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

Flier, B., F.A.J. Bosch, H.W. Volberda. 2003. Co-evolution in Strategic Renewal Behaviour of British, Dutch and French Financial Incumbents: Interaction of Environmental Selection, Institutional Effects and Managerial Intentionality. *Journal of Management Studies* **40**(8) 2163-2187.

Fosfuri, A. 2006. The licensing dilemma: understanding the determinants of the rate of technology licensing. *Strategic Management Journal* **27**(12) 1141-1158.

Gans, J.S., D.H. Hsu, S. Stern. 2008. The impact of uncertain intellectual property rights on the market for ideas: Evidence from patent grant delays. *Management Science* **54**(5) 982-997.

Garud, R., S. Jain, A. Kumaraswamy. 2002. Institutional entrepreneurship in the sponsorship of common technological standards: The case of Sun Microsystems and Java. *Academy of Management Journal* **45**(1) 196-214.

George, G., R. Kotha, Y. Zheng. 2008. Entry into insular domains: A longitudinal study of knowledge structuration and innovation in biotechnology firms. *Journal of Management Studies* **45**(8) 1448-1474.

Gittelman, M., B. Kogut. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science* **49**(4) 366-382.

Greene, W. 2007. *Functional form and heterogeneity in models for count data.* Now Publishers Inc.

Greenwood, R., C.R. Hinings. 1996. Understanding radical organizational change: Bringing together the old and the new institutionalism. *Academy of management review* 1022-1054.

Guimaraes, P. 2008. The fixed effects negative binomial model revisited. *Economics Letters* **99**(1) 63-66.

Gu, F. 2005. Innovation, future earnings, and market efficiency. *Journal of Accounting, Auditing & Finance* **20**(4) 385-418.

Gulati, R. 1995. Does familiarity breed trust? The implications of repeated ties for contractual choice in alliances. *Academy of management journal* **38**(1) 85-112.

Hall, B.H., A. Jaffe, M. Trajtenberg. 2005. Market value and patent citations. *RAND Journal of economics* 16-38.

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

Hannan, M.T., J. Freeman. 1993. *Organizational ecology*. Harvard University Press.

Harhoff, D., F.M. Scherer, K. Vopel. 2003. Citations, family size, opposition and the value of patent rights. *Research Policy* **32**(8) 1343-1363.

Harhoff, D., S. Wagner. 2006. Modeling the duration of patent examination at the European Patent Office. SFB/TR 15 Discussion Paper.

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

Hausman, J. A., B. H. Hall, Z. Griliches. 1984. Econometric models for count data with an application to the patents–R & D relationship. *Econometrica* **52** 909–938.

He, J., H.C. Wang. 2009. Innovative knowledge assets and economic performance: The asymmetric roles of incentives and monitoring. *Academy of Management Journal* **52**(5) 919-938.

Heeley, M.B., R. Jacobson. 2008. The recency of technological inputs and financial performance. *Strategic Management Journal* **29**(7) 723-744.

Heugens, P.P., M.W. Lander. 2009. Structure! Agency!(and other quarrels): A meta-analysis of institutional theories of organization. *Academy of Management Journal* **52**(1) 61-85.

Hodgson, G.M. 2004. *The evolution of institutional economics: agency, structure, and Darwinism in American institutionalism*. Psychology Press.

Hodgson, G. M. 2013. Understanding organizational evolution: Toward a research agenda using generalized Darwinism. *Organization Studies* **34**(7) 973-992.

Hoetker, G., R. Agarwal. 2007. DEATH HURTS, BUT IT ISN'T FATAL: THE POSTEXIT DIFFUSION OF KNOWLEDGE CREATED BY INNOVATIVE COMPANIES. *Academy of Management Journal* **50**(2) 446-467.

Hoffman, A.J. 1999. Institutional evolution and change: Environmentalism and the US chemical industry. *Academy of Management Journal* 351-371.

Huygens, M., C. Baden-Fuller, F.A.J. Van Den Bosch, H.W. Voiberda. 2001. Co-evolution of Firm Capabilities and Industry Competition: Investigating the Music Industry. *Organization Studies* **22**(6) 971-1011.

BIBLIOGRAPHY

Jain, S., G. George. 2007. Technology transfer offices as institutional entrepreneurs: the case of Wisconsin Alumni Research Foundation and human embryonic stem cells. *Industrial and Corporate Change* **16**(4) 535-567.

Joshi, A.M., A. Nerkar. 2011. When do strategic alliances inhibit innovation by firms? Evidence from patent pools in the global optical disc industry. Strategic Management Journal **32**(11) 1139-1160.

Kaplan, S., M. Tripsas. 2008. Thinking about technology: Applying a cognitive lens to technical change. *Research Policy* **37**(5) 790-805.

Khoury, T.A., E.G. Pleggenkuhle-Miles. 2011. Shared inventions and the evolution of capabilities: Examining the biotechnology industry. *Research Policy* **40**(7) 943-956.

Koellinger, P. 2008. The relationship between technology, innovation, and firm performance. Empirical evidence from e-business in Europe. *Research Policy* **37**(8) 1317-1328.

Kondo, M. 1999. R&D dynamics of creating patents in the Japanese industry. *Research Policy* **28**(6) 587-600.

Leone, M.I., T. Reichstein. 2012. Licensing-in fosters rapid invention! the effect of the grant-back clause and technological unfamiliarity. *Strategic Management Journal* **33**(8) 965-985.

Lettl, C., K. Rost, I. Von Wartburg. 2009. Why are some independent inventors 'heroes' and others 'hobbyists'? The moderating role of technological diversity and specialization. *Research Policy* **38**(2) 243-254.

Levit, G.S., U. Hossfeld, U. Witt. 2011. Can Darwinism be "Generalized" and of what use would this be? *Journal of Evolutionary Economics* **21**(4) 545-562.

Lewin, A.Y., H.W. Volberda. 1999. Prolegomena on coevolution: A framework for research on strategy and new organizational forms. *Organization science* **10**(5) 519-534.

Lewin, A.Y., C.B. Weigelt, J.D. Emery. 2004. Adaptation and selection in strategy and change. M.S. Poole, A.H. Van de Ven, eds. *Handbook of organizational change and innovation*. Oxford University Press, 108-160.

Markman, G.D., M.I. Espina, P.H. Phan. 2004. Patents as surrogates for inimitable and non-substitutable resources. *Journal of Management* **30**(4) 529-544.

Meyer, M. 2000. Does science push technology? Patents citing scientific literature. *Research Policy* **29**(3) 409-434.

Miller, D.J., M.J. Fern, L.B. Cardinal. 2007. The use of knowledge for technological innovation within diversified firms. *Academy of Management Journal* **50**(2) 307-325.

Mina, A., R. Ramlogan, G. Tampubolon, J.S. Metcalfe. 2007. Mapping evolutionary trajectories: Applications to the growth and transformation of medical knowledge. *Research Policy* **36**(5) 789-806.

Mowery, D.C., A.A. Ziedonis. 2002. Academic patent quality and quantity before and after the Bayh-Dole act in the United States. *Research Policy* **31**(3) 399-418.

Munir, K.A., M. Jones. 2004. Discontinuity and after: the social dynamics of technology evolution and dominance. *Organization Studies* **25**(4) 561-581.

Munir, K.A., N. Phillips. 2005. The birth of the ‘Kodak Moment’: institutional entrepreneurship and the adoption of new technologies. *Organization Studies* **26**(11) 1665-1687.

Murmann, J.P. 2013. The coevolution of industries and important features of their environments. Organization Science **24**(1) 58-78.

Nagaoka, S. 2007. Assessing the R&D management of a firm in terms of speed and science linkage: evidence from the US patents. *Journal of Economics & Management Strategy* **16**(1) 129-156.

Nelson, R.R., S.G. Winter. 1982. *An evolutionary theory of economic change*. Belknap press.

Nemet, G.F., E. Johnson. 2012. Do important inventions benefit from knowledge originating in other technological domains? *Research Policy* **41**(1) 190-200.

Nerkar, A. 2003. Old is gold? The value of temporal exploration in the creation of new knowledge. *Management Science* **49**(2) 211-229.

Nooteboom, B. 1992. Towards a dynamic theory of transactions. *Journal of evolutionary economics* **2**(4) 281-299.

Nooteboom, B. 1999. *Inter-firm alliances: Analysis and design*. Psychology Press.

Nooteboom, B. 2000. *Learning and innovation in organizations and economies*. OUP Oxford.

Nooteboom, B., W. Van Haverbeke, G. Duysters, V. Gilsing, A. Van den Oord. 2007. Optimal cognitive distance and absorptive capacity. *Research policy* **36**(7) 1016-1034.

BIBLIOGRAPHY

Operti, E., G. Carnabuci. 2014. Public Knowledge, Private Gain The Effect of Spillover Networks on Firms' Innovative Performance. *Journal of Management* **40**(4) 1042-1074.

Pacheco, D.F., J.G. York, T.J. Dean, S.D. Sarasvathy. 2010. The coevolution of institutional entrepreneurship: A tale of two theories. *Journal of Management* **36**(4) 974-1010.

Penner-Hahn, J., J.M. Shaver. 2005. Does international research and development increase patent output? An analysis of Japanese pharmaceutical firms. *Strategic Management Journal* **26**(2) 121-140.

Phene, A., K. Fladmoe-Lindquist, L. Marsh. 2006. Breakthrough innovations in the US biotechnology industry: the effects of technological space and geographic origin. *Strategic Management Journal* **27**(4) 369-388.

Powell, W.W., K.W. Koput, L. Smith-Doerr. 1996. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly* 116-145.

Rindova, V.P., A.P. Petkova. 2007. When is a new thing a good thing? Technological change, product form design, and perceptions of value for product innovations. *Organization science* **18**(2) 217-232.

Robinson, K.C., P.P. McDougall. 1998. The impact of alternative operationalizations of industry structural elements on measures of performance for entrepreneurial manufacturing ventures. *Strategic Management Journal* **19**(11) 1079-1100.

Rodrigues, S., J. Child. 2003. Co-evolution in an Institutionalized Environment. *Journal of Management Studies* **40**(8) 2137-2162.

Rosenkopf, L., A. Nerkar. 2001. Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry. *Strategic Management Journal* **22**(4) 287-306.

Rosenkopf, L., M.L. Tushman. 1994. The coevolution of technology and organization. J.A.C. Baum, J.V. Singh, eds. *Evolutionary dynamics of organizations*. Oxford University Press, 403-424.

Rothaermel, F.T., W. Boeker. 2008. Old technology meets new technology: complementarities, similarities, and alliance formation. *Strategic Management Journal* **29**(1) 47-77.

Rothaermel, F.T., A.M. Hess. 2007. Building dynamic capabilities: Innovation driven by individual-, firm-, and network-level effects. *Organization Science* **18**(6) 898-921.

Rothaermel, F.T., C.W.L. Hill. 2005. Technological discontinuities and complementary assets: A longitudinal study of industry and firm performance. *Organization Science* **16**(1) 52-70.

Sapsalis, E.r., B. van Pottelsberghe de la Potterie, R. Navon. 2006. Academic versus industry patenting: An in-depth analysis of what determines patent value. *Research Policy* **35**(10) 1631-1645.

Schoenmakers, W., G. Duysters. 2010. The technological origins of radical inventions. *Research Policy* **39**(8) 1051-1059.

Schuler, D.A., K. Rehbein, R.D. Cramer. 2002. PURSUING STRATEGIC ADVANTAGE THROUGH POLITICAL MEANS: A MULTIVARIATE APPROACH. *Academy of Management Journal* **45**(4) 659-672.

Shane, S. 2001. Technological Opportunities and New Firm Creation. *Management Science* **47**(2) 205-220.

Smircich, L. 1983. Organizations as shared meanings. *Organizational symbolism* **55** 65.

Somaya, D., I.O. Williamson, X. Zhang. 2007. Combining patent law expertise with R&D for patenting performance. *Organization Science* **18**(6) 922-937.

Sørensen, J.B., T.E. Stuart. 2000. Aging, obsolescence, and organizational innovation. *Administrative Science Quarterly* **45**(1) 81-112.

Sorenson, O., L. Fleming. 2004. Science and the diffusion of knowledge. *Research Policy* **33**(10) 1615-1634.

Srivastava, M.K., D.R. Gnyawali. 2011. When do relational resources matter? Leveraging portfolio technological resources for breakthrough innovation. *Academy of Management Journal* **54**(4) 797-810.

Stern, R.N., S.R. Barley. 1996. Organizations and social systems: Organization theory's neglected mandate. *Administrative Science Quarterly* 146-162.

Stuart, T.E. 1998. Network positions and propensities to collaborate: An investigation of strategic alliance formation in a high-technology industry. *Administrative Science Quarterly* 668-698.

Stuart, T.E. 1999. A structural perspective on organizational innovation. *Industrial and Corporate Change* **8**(4) 745-775.

Stuart, T.E. 2000. Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. *Strategic Management Journal* **21**(8) 791-811.

Stuart, T.E., J.M. Podolny. 1996. Local search and the evolution of technological capabilities. *Strategic Management Journal* **17**(S1) 21-38.

Teece, D.J. 2009. *Dynamic Capabilities and Strategic Management*. Oxford University Press.

The Thomson Corporation 2007. GLOBAL PATENT SOURCES. An Overview of International Patents, 6 ed. Thomson Scientific, London.

Trajtenberg, M., R. Henderson, A. Jaffe. 1997. University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and new technology* **5**(1) 19-50.

Tushman, M., J. Murmann. 1998. Dominant designs, technology cycles, and organizational outcomes. *Research in Organizational Behavior* 20 231–266.

Valentini, G. 2012. Measuring the effect of M&A on patenting quantity and quality. *Strategic Management Journal* **33**(3) 336-346.

Valentini, G., M. C. Di Guardo. 2012. M&A and the profile of inventive activity. *Strategic Organization* **10** 384-405.

Van de Ven, A.H., R. Garud. 1994. The coevolution of technical and institutional events in the development of an innovation. J.A.C. Baum, J.V. Singh, eds. *Evolutionary dynamics of organizations*. Oxford University Press, 425-443.

Van Den Bergh, J. C. J. M., S. Stagl. 2003. Coevolution of economic behaviour and institutions: towards a theory of institutional change. *Journal of evolutionary economics* **13**(3) 289-317.

Van Zeebroeck, N., B. Van Pottelsberghe de la Potterie, D. Guellec. 2009. Claiming more: the increased voluminosity of patent applications and its determinants. *Research Policy* **38**(6) 1006-1020.

Vanhaverbeke, W., V. Gilsing, B. Beerkens, G. Duysters. 2009. The Role of Alliance Network Redundancy in the Creation of Core and Non-core Technologies. *Journal of Management Studies* **46**(2) 215-244.

Volberda, H.W., A.Y. Lewin. 2003. Co-evolutionary Dynamics Within and Between Firms: From Evolution to Co-evolution. *Journal of Management Studies* **40**(8) 2111-2136.

Weick, K. 1969. The social psychology of organizing. *Reading, Addison-Wesley*.

Weick, K. 1995. *Sensemaking in organizations*. Sage.

Wohlgezogen, F., P. Hirsch. 2009. Negotiating actor-environment relations: A framework for comparative research. *Research in the Sociology of Organizations* **26** 153-182.

Wooldridge, J.M. 2002. Introductory Econometrics. *South-Western, Mason, Ohio*.

Yayavaram, S., G. Ahuja. 2008. Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability. *Administrative Science Quarterly* **53**(2) 333-362.

APPENDIX A: Sample Firms

BIOTECHNOLOGY FIRMS

Name	Country
1. Affymetrix	USA
2. Alizyme	UK
3. Amgen	USA
4. Axis Shield	UK
5. Biogen	USA
6. Biotie Therapies	Finland
7. BTG	UK
8. Celgene	USA
9. Crucell	The Netherlands
10. Epigenomics	Germany
11. Genmab	Denmark
12. Genus	UK
13. Genzyme	USA
14. Gilead Sciences	USA
15. Human Genome Sciences	USA
16. Incyte	USA
17. Intercell	Austria
18. InterMune	USA
19. Karo Bio	Sweden
20. Kyowa Hakko	Japan
21. MediGene	Germany
22. Merial	UK
23. Morphosys	Germany
24. Novozymes	Denmark

Name	Country
25. OSI	USA
26. Qiagen	The Netherlands
27. Transgene	France
28. Vernalis	UK
29. VERTEX PHARM	USA

PHARMACEUTICAL FIRMS

Name	Country
1. Abbott	France
2. Allergan	UK
3. AstraZeneca	UK
4. Biotest	Germany
5. Boehringer Ingelheim	Denmark
6. Bristol-Myers Squibb	Belgium
7. Cephalon	Denmark
8. Cerep	UK
9. Egis Pharmaceuticals	Ireland
10. Eisai	Luxembourg
11. Elan	Slovenia
12. Eli Lilly	Hungary
13. Evotec	Italy
14. Exelixis	Germany
15. Forest	Spain
16. Gedeon Richter	Germany
17. GlaxoSmithKline	France
18. Guerbet	UK
19. GW Pharmaceuticals	Hungary
20. Ipsen	France

APPENDIX A: SAMPLE FIRMS

Name	Country
21. Johnson	Germany
22. Kissei	UK
23. Krka	UK
24. Lundbeck	France
25. Mylan	Switzerland
26. NicOx	USA
27. Nippon	Switzerland
28. Novartis	USA
29. Novo Nordisk	USA
30. Ono	USA
31. Oxford Biomedica	USA
32. Pfizer	Japan
33. Recordati	USA
34. Roche	Israel
35. Sanofi-Aventis	USA
36. Santen	Japan
37. Sepracor	Japan
38. Shionogi	USA
39. Shire	Japan
40. Stada Arzneimittel	USA
41. Taisho	USA
42. Teva	USA
43. UCB	Japan
44. Vectura	USA
45. Watson	Japan
46. Zeltia	Japan

CHEMICAL FIRMS

APPENDIX A: SAMPLE FIRMS

Name	**Country**
1. Ahlstrom	Finland
2. Air Liquide	France
3. Air Products and Chemicals	USA
4. AKZO Nobel	The Netherlands
5. ALTANA	Germany
6. Asahi Kasei	Japan
7. Auriga Industries	Denmark
8. Avery Dennison	USA
9. BASF	Germany
10. Bayer	Germany
11. Borealis	Denmark
12. Celanese	Germany
13. Clariant	Switzerland
14. Cognis	Germany
15. Croda International	UK
16. Daicel Chemical Industries	Japan
17. Denki Kagaku Kogyo	Japan
18. Dow Chemical	USA
19. DSM	The Netherlands
20. Dynaction	France
21. Eastman Chemical	USA
22. FMC	USA
23. Fuchs Petrolub	Germany
24. Johnson Matthey	UK
25. JSR	Japan
26. K+S	Germany
27. Kaneka	Japan
28. Kemira	Finland
29. Kuraray	Japan

APPENDIX A: SAMPLE FIRMS

Name	Country
30. Linde	Germany
31. Lonza	Switzerland
32. Lubrizol	USA
33. Mitsubishi Chemical	Japan
34. Mitsubishi Gas Chemical	Japan
35. Mitsubishi Rayon	Japan
36. Mitsui Chemicals	Japan
37. Nippon Kayaku	Japan
38. Nippon Shokubai	Japan
39. Nitto Denko	Japan
40. Nolato	Sweden
41. PPG Industries	USA
42. Praxair	USA
43. Recticel	Belgium
44. Rhodia	France
45. SGL Carbon	Germany
46. Shin-Etsu Chemical	Japan
47. Showa Denko	Japan
48. SNPE	France
49. Solvay	Belgium
50. Sud-Chemie	Germany
51. Sumitomo Bakelite	Japan
52. Sumitomo Chemical	Japan
53. Syngenta	Switzerland
54. Teijin	Japan
55. Tessenderlo	Belgium
56. Tokuyama	Japan
57. Toray Industries	Japan
58. Tosoh	Japan

Name	**Country**
59. Toyobo	Japan
60. UBE Industries	Japan
61. Umicore	Belgium
62. Wacker-Chemie	Germany
63. Yule Catto	UK
64. Zeon	Japan

APPENDIX B: A Complete Patent Family

Derwent Innovation Index. A complete patent family of the firm NOVOZYMES applied in 2004.	
Title	Preparation of feed composition for feeding animals comprises hydrolyzing fish meat with neutral protease and alkaline protease, and inactivating proteases through heat treatment
Patent Numbers	WO2004016098-A1; AU2003250319-A1; NO200501267-A; EP1531682-A1; BR200313393-A; JP2005535337-W; CN1674790-A; KR2005036969-A; US2006204612-A1; NZ538120-A; AU2003250319-B2; CN100435648-C; EP1531682-B1; DE60335113-E; JP4727989-B2
Inventors	PEDERSEN B P; STANDAL H; PEDERSEN P B; HAAKON S; PIIL P B
Patent Assignees and Codes	NOVOZYMES AS (NOVO); DENOFA AS (DENO-Non-standard); ORAVEIEN INDUSTRIPARK AS (ORAV-Non-standard)
Derwent Primary Accession Number	2004-203703

APPENDIX B: A COMPLETE PATENT FAMILY

<table>
<tr>
<td>Patents Cited by Inventor / Examiner</td>
<td>WO2004016098-A1 –
EP301795-A ASAHI DENKA KOGYO KK (ASAE) SHIRAKAWA Y, MINOWA Y, AZUMI T, HISANO J;
US3578461-A MONSANTO CO (MONS);
US3697285-A ROHM & HAAS CO (ROHM);
US3924005-A SOC PROD NESTLE SA (NEST);
US4036993-A TENSEI SUISAN CO (TENS-Non-standard);
US4473589-A FREEMAN L D (FREE-Individual) FREEMAN L D, SAWHILL J W;

EP1531682-B1 –
EP301795-A ASAHI DENKA KOGYO KK (ASAE) SHIRAKAWA Y, MINOWA Y, AZUMI T, HISANO J;
US3578461-A MONSANTO CO (MONS);
US3697285-A ROHM & HAAS CO (ROHM);
US3924005-A SOC PROD NESTLE SA (NEST);
US4036993-A TENSEI SUISAN CO (TENS-Non-standard);
US4473589-A FREEMAN L D (FREE-Individual) FREEMAN L D, SAWHILL J W</td>
</tr>
<tr>
<td>Citing Patents</td>
<td>
<table>
<tr><td>Derwent Primary Accession Number</td><td>Title</td></tr>
<tr><td>2006-586470</td><td>New anti-hypertensive Salmo or Oncorhynchus protein hydrolysate, useful as a dietary supplement for reducing mean blood pressure and for treating or preventing hypertension</td></tr>
<tr><td>2010-J71856</td><td>Preparation of fish meal useful as seasonings and health food, by slicing viscera removed fish, pulverizing fish, preparing first and second decomposition material, mixing second decomposition material with pulverized fish</td></tr>
<tr><td>2009-R40544</td><td>Method of obtaining fodder based on protein hydrolysate</td></tr>
</table>
</td>
</tr>
<tr>
<td>Articles Cited by Inventor / Examiner</td>
<td>LIASET ET AL: "Studies on the nitrogen recovery in enzymatic hydrolysis of Atlantic salmon frames by Protamex protease" PROCESS BIOCHEMISTRY, vol. 37, 2000, pages 1263-1269.</td>
</tr>
</table>

APPENDIX B: A COMPLETE PATENT FAMILY

Abstract	NOVELTY - Preparing a feed composition for feeding animals comprising hydrolyzing fish meat with neutral protease and alkaline protease, and inactivating the proteases through heat treatment, is new. USE - The method is useful for preparing a feed composition for feeding animals (claimed). ADVANTAGE - The invention achieves higher growth rate of the animals.
International Patent Classification (IPC)	A23K-001/10; A23K-001/165; A23G-004/00; A23J-003/34; A23K-001/18
Derwent Class Codes	B04 (Natural products and polymers, testing, compounds of unknown structure); C03 (Other organic or inorganic compounds and multi-component mixtures); D12 (Butchering, meat treatment, processing poultry or fish); D13 (Other foodstuffs and treatment); D16 (Fermentation industry)
Derwent Manual Codes	B04-B04M; B04-L05C; C04-B04M; C04-L05C; D02-A03A; D03-G05; D05-A02C

APPENDIX B: A COMPLETE PATENT FAMILY

Patents Details	*Patent Number*	*Patent Publication Date*	*Main IPC*	*Derwent Week*	Page Count	Language
	WO2004016098-A1	26 Feb 2004	A23K-001/10	200419	Pages: 19	English
	AU2003250319-A1	03 Mar 2004		200457		
	NO200501267-A	11 Mar 2005		200530		
	EP1531682-A1	11 Mar 2005		200535		English
	BR200313393-A	21 Jun 2005		200542		
	JP2005535337-W	24 Nov 2005	A23K-001/10	200581	Pages: 15	
	CN1674790-A	28 Sep 2005		200610		
	KR20050036969-A	20 Apr 2005 200643		200643		
	US2006204612-A1	14 Sep 2006	A23G-004/00	200661		
	NZ538120-A	29 Sep 2006	A23G-004/00	200661		
	AU2003250319-B2	05 Jun 2008	A23K-001/10	200862		
	CN100435648-C	26 Nov 2008	A23K-001/10	200936		Chinese
	EP1531682-B1	24 Nov 2010	A23K-001/10	201077		English
	DE60335113-E	05 Jan 2011	A23K-001/10	201105		German
	JP4727989-B2	20 Jul 2011	A23K-001/10	201148	Pages: 12	Japanese
Application Details and Date	WO2004016098-A1	WODK00535	12 Aug 2003			
	AU2003250319-A1	AU250319	12 Aug 2003			
	NO200501267-A	NO001267	11 Mar 2005			
	EP1531682-A1	EP787741	12 Aug 2003			

APPENDIX B: A COMPLETE PATENT FAMILY

	BR200313393-A	BR013393	12 Aug 2003
	JP2005535337-W	JP528440	12 Aug 2003
	CN1674790-A	CN819199	12 Aug 2003
	KR2005036969-A	KR702447	14 Feb 2005
	US2006204612-A1	US524934	21 Feb 2006
	NZ538120-A	NZ538120	12 Aug 2003
	AU2003250319-B2	AU250319	12 Aug 2003
	CN100435648-C	CN819199	12 Aug 2003
	EP1531682-B1	EP787741	12 Aug 2003
	DE60335113-E	DE635113	12 Aug 2003
	JP4727989-B2	JP528440	12 Aug 2003
Priority Application Information and Date	DK001207		14 Aug 2002
	DK000762		20 May 2003
Designated States	*WO2004016098-A1*: (National): AE; AG; AL; AM; AT; AU; AZ; BA; BB; BG; BR; BY; BZ; CA; CH; CN; CO; CR; CU; CZ; DE; DK; DM; DZ; EC; EE; ES; FI; GB; GD; GE; GH; GM; HR; HU; ID; IL; IN; IS; JP; KE; KG; KP; KR; KZ; LC; LK; LR; LS; LT; LU; LV; MA; MD; MG; MK; MN; MW; MX; MZ; NI; NO; NZ; OM; PG; PH; PL; PT; RO; RU; SC; SD; SE; SG; SK; SL; SY; TJ; TM; TN; TR; TT; TZ; UA; UG; US; UZ; VC; VN; YU; ZA; ZM; ZW (Regional): AT; BE; BG; CH; CY; CZ; DE; DK; EA; EE; ES; FI; FR; GB; GH; GM; GR; HU; IE; IT; KE; LS; LU; MC; MW; MZ; NL; OA; PT; RO; SD; SE; SI; SK; SL; SZ; TR; TZ; UG; ZM; ZW; *EP1531682-A1*: (Regional): AL; AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR; HU; IE; IT; LI; LT; LU; LV; MC; MK; NL; PT; RO; SE; SI; SK; TR; EP1531682-B1: (Regional): AT; BE; BG; CH; CY; CZ; DE; DK; EE; ES; FI; FR; GB; GR; HU; IE; IT; LI; LU; MC; NL; PT; RO; SE; SI; SK; TR		

APPENDIX B: A COMPLETE PATENT FAMILY

Further Application Details	AU2003250319-A1 Based on Patent WO2004016098; NO200501267-A PCT application Application WODK00535; EP1531682-A1 Based on Patent WO2004016098; EP1531682-A1 PCT application Application WODK00535; BR200313393-A Based on Patent WO2004016098; BR200313393-A PCT application Application WODK00535; JP2005535337-W Based on Patent WO2004016098; JP2005535337-W PCT application Application WODK00535; KR2005036969-A Based on Patent WO2004016098; KR2005036969-A PCT application Application WODK00535; US2006204612-A1 PCT application Application WODK00535; NZ538120-A Based on Patent WO2004016098; NZ538120-A PCT application Application WODK00535; AU2003250319-B2 Based on Patent WO2004016098; EP1531682-B1 PCT application Application WODK00535; EP1531682-B1 Based on Patent WO2004016098; DE60335113-E PCT application Application WODK00535; DE60335113-E EP application Application EP787741; DE60335113-E Based on Patent WO2004016098; DE60335113-E Based on Patent EP1531682; JP4727989-B2 PCT application Application WODK00535; JP4727989-B2 Based on Patent WO2004016098; JP4727989-B2 Previous Publ. Patent JP2005535337
Field of Search	x

APPENDIX C: MySQL Tables

Table Name: ***Main_Patents***

Description: Contains the full information of the patent families that have been applied by the firms of our sample for the period from 2003 to 2009

Table Type: InnoDB

Field Name	*Type*	*Extra*
Id_system	Integer	Primary Key Auto Increment
Company Code	Integer	
Company Name	Text	
Field	Text	
Title	Text	
Patent Numbers	Text	
Inventors	Text	
Patent Assignees and Codes	Text	
Derwent Primary Accession Number	Text	
Articles Cited by Inventor / Examiner	Text	
Abstract	Text	
International Patent Classification (IPC)	Text	
Derwent Class Codes	Text	
Derwent Manual Codes	Text	
Patents Details	Text	
Application Details and Date	Text	
Priority Application Information and Date	Text	
Designated States	Text	
Further Application Details	Text	
Field of Search	Text	

APPENDIX C: MySQL Tables

Table Name: ***Forward_Patents***

Description: Contains the full information of the patent families that have cited the patent families of the table Main_Patents until December 2011

Table Type: InnoDB

The fields of the table *Forward_Patents* are the same as those of the table *Main_Patents.*

Table Name: ***Main_Forward***

Description: Connects each patent family of the table *Main_Patents* with the patent families of the table *Forward_Patents*

Table Type: InnoDB

Field Name	*Type*	*Extra*
Id_system	Integer	Primary Key Auto Increment
Main Derwent Primary Accession Number	Text	
Forward Derwent Primary Accession Number	Text	

Table Name: ***Backward_Patents***

Description: Contains the full information of the patent families that have been cited by the patent families of the table *Main_Patents*.

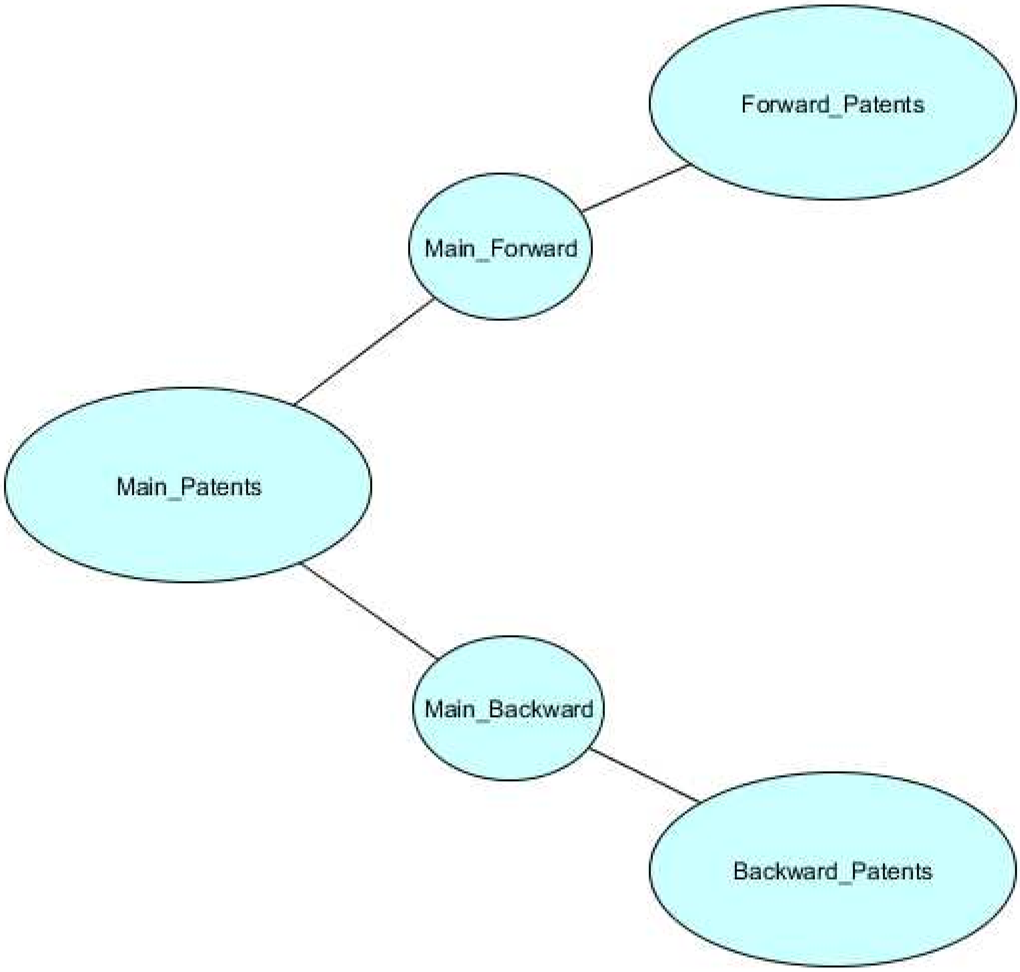
Table Type: InnoDB

The fields of the table *Backward_Patents* are the same as those of the table *Main_Patents.*

APPENDIX C: MySQL Tables

Table Name: ***Main_Backward*** *Description*: Connects each patent family of the table *Main_Patents* with the patent families of the table *Backward_Patents* *Table Type:* InnoDB		
Field Name	*Type*	*Extra*
Id_system	Integer	Primary Key Auto Increment
Main Derwent Primary Accession Number	Text	
Backward Derwent Primary Accession Number	Text	

Figure Appendix C.1 The Schema of our Database



APPENDIX D: Patent-based Variables' Description in *SQL Query Style*

Patent-based Variable	*Description in SQL Query Style*	*Result*
Technological Impact	select distinct *Company Name*, *Company Code*, count (*Main_Forward.Id_system*) from *Main_Patents*, *Forward_Main* where *Priority Application Information and Date* like '2003%' and *Main_Patents.Derwent Primary Accession Number*=*Main_Forward.Main Derwent Primary Accession Number* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the year of the first application of the patent family based on the field Priority Application Information and Date of the table Main_Patents.*	Our sample's firms and the number of citations that their patent families applied in year 2003 have received.
Breadth of Impact – *Beyond-Industry Impact*	select distinct *Company Name*, *Company Code*, count (*Main_Forward.Id_system*) from *Main_Patents*, *Forward_Main*, *Forward_Patents* where *Priority Application Information and Date like '2003%'* and *Main_Patents.Derwent Primary Accession Number*=*Main_Forward.Main Derwent Primary Accession Number* and *Forward_Patents.Derwent Primary Accession Number*=*Main_Forward. Forward Derwent Primary Accession Number* and *Forward_Patents.International Patent Classification* NOT IN Industry_List_of_IPC* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the Industry_List_of_IPC based on the field International Patent Classification of the table Main_Patents of all the patent families of the firms that belong to the focal firm's industry in the years 2001-2003.*	Our sample's firms and the number of citations that their patent families applied in year 2003 have received from patent families that are classified in international patent classes which are different from the total of all international patent classes of the focal firm's industry

APPENDIX D: Patent-based Variables' Description in SQL Query Style

Patent-based Variable	*Description in SQL Query Style*	*Result*
Breadth of Impact – *Beyond-Firm Impact*	select distinct *Company Name*, *Company Code*, count (*Main Forward.Id_system*) from *Main_Patents*, *Forward_Main*, *Forward_Patents* where *Priority Application Information and Date like '2003%'* and *Main_Patents.Derwent Primary Accession Number=Main_Forward.Main Derwent Primary Accession Number* and *Forward_Patents.Derwent Primary Accession Number=Main_Forward. Forward Derwent Primary Accession Number* and *Forward_Patents.International Patent Classification* NOT IN Firm_List_of_IPC* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the Firm_List_of_IPC based on the field International Patent Classification of the table Main_Patents of all the patent families of the focal firm in the years 2001-2003.*	Our sample's firms and the number of citations that their patent families applied in year 2003 have received from patent families that are classified in international patent classes which are different from the total of all international patent classes of the focal firm

APPENDIX D: Patent-based Variables' Description in SQL Query Style

Patent-based Variable	*Description in SQL Query Style*	*Result*

Breadth of Impact – *Herfindahl Type Breadth*	*FOR EACH* Company Code *DO*: {
	FOR EACH Patent Family in 2003 *DO*: {
	select distinct *Forward_Patents.International Patent Classification*, count (*Main Forward.Id_system*) from *Main_Patents, Forward_Main, Forward_Patents* where *Priority Application Information and Date like '2003%'* and *Main_Patents.Derwent Primary Accession Number=Main_Forward. Main Derwent Primary Accession Number* and *Forward_Patents.Derwent Primary Accession Number=Main_Forward. Forward Derwent Primary Accession Number* group by *Forward_Patents.International Patent Classification* order by *Forward_Patents.International Patent Classification* *(returns the number of forward citations per International Patent Classification)*
	}}.*
	
	**A special script in PHP was used to build the firm's average Herfindahl score based on the Herfindahl score of each patent family.*
		Our sample's firms and the average score of *Herfindahl Type Breadth* of their patent families applied in year 2003

Scientific Basis – *Non Patent References*	select distinct *Company Name*, *Company Code*, average (*Articles Cited by Inventor / Examiner**) from *Main_Patents* where *Priority Application Information and Date like '2003%'* group by *Company Code* order by *Company Code* **A special script in PHP was used to build the number of Articles Cited by Inventor / Examiner from the field Articles Cited by Inventor / Examiner of the table Main_Patents.*	Our sample's firms and the average number of *Non Patent References* of their patent families applied in year 2003

APPENDIX D: Patent-based Variables' Description in SQL Query Style

Patent-based Variable	*Description in SQL Query Style*	*Result*
Scientific Basis – *Scientific Inventions*	select distinct *Company Name*, *Company Code*, count (*Id_system*) from *Main_Patents* where *Priority Application Information and Date like '2003%'* and *Patent Assignees and Codes** like '%UNIV%' group by *Company Code* order by *Company Code* **A special script in PHP was used to distinguish Patent Assignees Names from Patent Assignees Codes in the field Patent Assignees and Codes of the table Main_Patents.*	Our sample's firms and the number of their patent families that have been applied by the firm in cooperation with a university applied in year 2003
Technological Collaboration	select distinct *Company Name*, *Company Code*, average (*Patent Assignees and Codes**) from *Main_Patents* where *Priority Application Information and Date like '2003%'* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the number of Patent Assignees and Codes from the field Patent Assignees and Codes of the table Main_Patents.*	Our sample's firms and the average number of the co-assignees of their patent families applied in 2003.
Technological Diversity – *IPC Diversity*	select distinct *Company Name*, *Company Code*, average (*International Patent Classification* *) from *Main_Patents* where *Priority Application Information and Date like '2003%'* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the number of International Patent Classifications of each patent family from the field International Patent Classification of the table Main_Patents.*	Our sample's firms and the average number of the International Patent Classifications of their patent families applied in 2003.

APPENDIX D: Patent-based Variables' Description in SQL Query Style

<u>*Patent-based Variable*</u>	<u>*Description in SQL Query Style*</u>	<u>*Result*</u>

Technological Diversity – *Originality*	*FOR EACH* Company Code *DO*: {
	FOR EACH Patent Family in 2003 *DO*: {
	select distinct *Backward_Patents.International Patent Classification*, count (*Main Backward.Id_system*) from *Main_Patents, Backward_Main*, *Backward_Patents* where *Main_Patents.Derwent Primary Accession Number= Main_Backward. Main Derwent Primary Accession Number* and *Backward_Patents.Derwent Primary Accession Number= Main_Backward. Backward Derwent Primary Accession Number* group by *Backward_Patents.International Patent Classification* order by *Backward_Patents.International Patent Classification* *(returns the number of backward citations per <u>International Patent Classification</u>)*
	}}.*
	
	**A special script in PHP was used to build the firm's average Herfindahl score based on the Herfindahl score of each patent family.*
		Our sample's firms and the average score of *Originality* of their patent families applied in year 2003

New-to-Firm Technological Domains	select distinct *Company Name*, *Company Code*, count (*Id_system*) from *Main_Patents* where *Priority Application Information and Date like '2003%'* and International Patent Classification NOT IN *List_of_4_years_IPC** group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the List_of_4_years_IPC based on the field <u>International Patent Classification</u> of the table <u>Main_Patents</u> of all the patent families of the focal firm in the previous 4 years.*	Our sample's firms and the number of their patent families that include at least one IPC that is new to the firm and are applied in year 2003

APPENDIX D: Patent-based Variables' Description in SQL Query Style

Patent-based Variable	*Description in SQL Query Style*	*Result*
Unexplored Technological Areas	select distinct *Company Name*, *Company Code*, count (*Main Backward.Id_system*) from *Main_Patents, Backward_Main* where *Priority Application Information and Date like '2003%'* and *Main_Patents.Derwent Primary Accession Number= Main_Backward.Main Derwent Primary Accession Number* group by *Company Code* order by *Company Code*	Our sample's firms and the number of citations that their patent families applied in year 2003 have made to prior patent families.
Recency of Technological Inputs	select distinct *Company Name*, *Company Code*, average (*Backward_Patents.Derwent Primary Accession Number**) from *Main_Patents, Backward_Main, Backward_Patents* where *Priority Application Information and Date like '2003%'* and *Main_Patents.Derwent Primary Accession Number= Main_Backward.Main Derwent Primary Accession Number* and *Backward_Patents.Derwent Primary Accession Number= Main_Backward. Backward Derwent Primary Accession Number* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the difference in years between the year that is included in the field Derwent Primary Accession Number of the table Main_Patents and the year that is included in the field Priority Application Information and Date of the table Backward_Patents*	Our sample's firms and the average time difference between the year of their patent families applied in 2003 and the year of application of the patent families that are cited by the focal patent families.

APPENDIX D: Patent-based Variables' Description in SQL Query Style

Patent-based Variable	*Description in SQL Query Style*	*Result*
Recent Technological Activity	select distinct *Company Name*, *Company Code*, count (*Id_system*) from *Main Patents* where *Priority Application Information and Date* >= '1998' and *Priority Application Information and Date* <= '2002' group by *Company Code* order by *Company Code* **A special script in PHP was used to get the number of year of application in numerical form from the field Priority Application Information and Date of the table Main_Patents.*	Our sample's firms and the number of the patent families each firm had applied for from 1998 until 2002.
Internal Focus	1. select distinct *Company Name*, *Company Code*, count (*Id_system*) from *Main Patents* where *Priority Application Information and Date like '2003%'* and *Patent Assignees and Codes* like '% *Company Code* %' group by *Company Code* order by *Company Code* *(returns the number of self-citations of the patent families applied in 2003 for each firm)* 2. select distinct *Company Name*, *Company Code*, count (*Main Backward.Id_system*) from *Main_Patents, Backward_Main* where *Priority Application Information and Date like '2003%'* and *Main_Patents.Derwent Primary Accession Number= Main_Backward.Main Derwent Primary Accession Number* group by *Company Code* order by *Company Code* *(returns the total number of backward citations of the patent families applied in 2003 for each firm)* **A special script in PHP was used to compute the ratio of the result of the first query (number of self-citations) to the result of the second query (number of backward citations)*	Our sample's firms and the average ratio of the citations that each patent family applied in 2003 makes to prior patent families held by the focal firm to its total backward citations

APPENDIX D: Patent-based Variables' Description in SQL Query Style

Patent-based Variable	*Description in SQL Query Style*	*Result*
Number of Claims	select distinct *Company Name*, *Company Code*, average (*number of claims**) from *Main_Patents* where *Priority Application Information and Date like '2003%'* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the number of claims from the field Abstract of table Main_Patents.*	Our sample's firms and the average number of the claims of their patent families applied in 2003.
Number of Inventors	select distinct *Company Name*, *Company Code*, average (*number of inventors**) from *Main_Patents* where *Priority Application Information and Date like '2003%'* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the number of inventors from the field Inventors of the table Main_Patents.*	Our sample's firms and the average number of inventors of their patent families applied in 2003.
Geographical Coverage – Patent Family Size	select distinct *Company Name*, *Company Code*, average (*number of patents**) from *Main_Patents* where *Priority Application Information and Date like '2003%'* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the number of patents from the field Patent Numbers of the table Main_Patents.*	Our sample's firms and the average number of patents that each patent family applied in 2003 includes.
Geographical Coverage – Designated States	select distinct *Company Name*, *Company Code*, average (*number of designated states **) from *Main_Patents* where *Priority Application Information and Date like '2003%'* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute the number of designated states from the field Designated States of the table Main_Patents.*	Our sample's firms and the average number of designated states of their patent families applied in 2003.

APPENDIX D: Patent-based Variables' Description in SQL Query Style

Patent-based Variable	*Description in SQL Query Style*	*Result*
Number of Patents	select distinct *Company Name*, *Company Code*, count (*Id_system*) from *Main Patents* where *Priority Application Information and Date like '2003%'* group by *Company Code* order by *Company Code*	Our sample's firms and the number of their patents applied in year 2003
Number of Granted Patents	select distinct *Company Name*, *Company Code*, count (*granted patents**) from *Main_Patents* where *Priority Application Information and Date like '2003%'* group by *Company Code* order by *Company Code* **A special script in PHP was used to compute if a patent family includes at least one patent that has been granted by a patent office. A special symbol at the end of each patent number (e.g. A1 or A) denotes if this patent has been granted or not. This information is included in the field Patent Numbers of the table Main_Patents.*	Our sample's firms and the average number of patent families applied in year 2003 that include at least one granted patent.

APPENDIX E: VIF Tests

Variable	**VIF**
1. Impact On Technological Evolution	3.22
2. Breadth of Impact – *Beyond-Industry Impact*	1.99
3. Breadth of Impact – *Beyond-Firm Impact*	3.34
4. Breadth of Impact – *Herfindahl Type Breadth*	3.02
5. Scientific Basis – *Non Patent References*	2.29
6. Scientific Basis – *Scientific Inventions*	1.07
7. Technological Collaboration	1.81
8. Technological Diversity – *IPC Diversity*	1.79
9. Technological Diversity – *Originality*	2.93
10. New-to-Firm Technological Domains	2.85
11. Unexplored Technological Areas	2.48
12. Recency of Technological Inputs	1.59
*13. Recent Technological Activity**	11.81
14. Internal Focus	1.38
*15. R&D Investment**	4.74
*16. Profitability**	1.39
*17. Firm Size**	5.23
18. Number of Claims	1.53
19. Number of Inventors	1.64
20. Geographical Coverage – *Patent Family Size*	3.49
21. Geographical Coverage – *Designated States **	2.04
*22. Number of Patents**	15.39

APPENDIX E: VIF Tests

23. Number of Granted Patents	2.86
Mean VIF	*3.47*
*Variables are logged.	

APPENDIX F: A FIXED-EFFECT NEGATIVE BINOMIAL MODEL FOR THE ABOVE THE AVERAGE INFLUENTIAL FIRMS

APPENDIX F: A Fixed-effect Negative Binomial Model Based on a Sample that Includes only the Firm-year Observations whose Patent Citations Count is Equal or Above the Average Value of a Specific Industry in a Specific Year.

Variables	**Patent citations count >= Average (Year, Industry)**
Scientific Basis – *Non Patent References*	-0.018 (0.048)
Scientific Basis – *Scientific Inventions*	-0.043 (0.075)
Technological Collaboration	-0.103+ (0.062)
Technological Collaboration ***Squared***	0.046+ (0.024)
Technological Diversity – *IPC Diversity*	0.007 (0.095)
Technological Diversity – *Originality*	-0.317**** (0.059)
New-to-Firm Technological Domains	-0.003 (0.033)
New-to-Firm Technological Domains ***Squared***	-0.005 (0.009)
Unexplored Technological Areas	0.068** (0.027)
Recency of Technological Inputs	0.127* (0.059)
Recent Technological Activity	-0.165

APPENDIX F: A FIXED-EFFECT NEGATIVE BINOMIAL MODEL FOR THE ABOVE THE AVERAGE INFLUENTIAL FIRMS

Variables	**<u>Patent citations count >= Average (Year, Industry)</u>**
	(0.279)
Internal Focus	-0.004 (0.052)
R&D Investment	-0.031 (0.187)
Lagged R&D Investment	0.548*** (0.157)
Firm Size	-0.193 (0.274)
Number of Patents	1.769**** (0.123)
Number of Claims	-0.006 (0.063)
Number of Inventors	0.110+ (0.065)
Geographical Coverage – *Patent Family Size*	0.188** (0.074)
Geographical Coverage – *Designated States*	0.263* (0.098)
Number of Granted Patents	0.273**** (0.061)
Profitability	0.181*** (0.057)
Lagged Profitability	0.007+ (0.004)

APPENDIX F: A FIXED-EFFECT NEGATIVE BINOMIAL MODEL FOR THE ABOVE THE AVERAGE INFLUENTIAL FIRMS

Variables	**Patent citations count >= Average (Year, Industry)**
USA (EU omitted)	0.401 (0.540)
JAPAN (EU omitted)	-0.083 (0.430)
Pharmaceuticals (Biotechnology omitted)	-0.469 (0.571)
Chemicals (Biotechnology omitted)	-1.065+ (0.591)
Years Dummies Included	
_cons	0.016 (0.514)
Log likelihood	-733.29
N	221
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001	