UNIVERSITA' DEGLI STUDI DI SALERNO

DIPARTIMENTO DI SCIENZE
ECONOMICHE E STATISTICHE

Dottorato di Ricerca in
Ingegneria ed Economia dell'Innovazione
XIV° ciclo

ANNO ACCADEMICO 2014/2015

TESI DI DOTTORATO IN

Five dimensions of innovation: Studying technological strategies through patent data

DOTTORANDO
ANTONELLO CAMMARANO

TUTOR
CH.MO PROF. MAURO CAPUTO

COORDINATORE DEL DOTTORATO
CH.MO PROF. ALESSANDRA AMENDOLA

CONTENTS

APPENDICES

A. PATSTAT database, data collection and example of framework application
B. List of companies

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This chapter clarifies the motivation for designing an integrated patent-based framework to evaluate innovation strategies and specifies the research question my PhD thesis seeks to answer. It also presents the scope, the research objectives and the structure of the thesis.

1.1 Premise

Technological innovation is the key driver for technological progress and firms' economic growth. Since firms pursue different innovation approaches, they achieve different innovation performances. Actually, a firm's ability to develop technologies and products is strongly conditioned by its stock of knowledge, expertise and technology from prior R&D (Arts, 2012). The R&D conducted by companies is an investment activity whose output is the firm's knowledge stock. This asset positively contributes to the firms' future financial performance and, then, to its market value (Hall *et al.*, 2005). The higher the level of innovativeness of the invention, the higher the expected technological and financial impact. As a matter of fact, some new technologies can be considered as an extension of previous innovations, whilst others are breakthrough, discontinuous or disruptive.

Analysing technological innovation requires objective and standardized data, thus scholars often refer to patents (Griliches, 1990). Actually, patents are a direct outcome of the inventive process, and more specifically of those inventions that are expected to have a commercial impact; furthermore, they capture the proprietary and competitive dimension of technological change (Archibugi and Pianta, 1996). Since obtaining patent protection is costly and time-consuming, only inventions that are expected to provide benefits that outweigh cost are applied. Patents have been treated as the most important output indicators of innovative activities and patent data have become the focus of many tools and techniques to measure innovation (Ma and Lee, 2008). Among the information available in such documents, technology classification, assignee field, citations and patent families are used to define different innovation metrics (Ahuja and Lampert, 2001; Graff, 2003; Belderbos *et al.*, 2010; Al-Ashaab *et al.*, 2011; Geum *et al.*, 2013).

1.2 Problem description and research question

The aim of this thesis is to define which strategical, technical and organizational issues affect innovative processes, their outputs and the quality of such innovation outputs. Actually, despite technological strategies implemented by innovative firms are widely investigated in literature, most attention has been devoted to only one dimension of R&D processes at time. Since innovation processes are featured by extreme complexity, I suggest a multidimensional approach, which provides a more complete overview of such processes, proposing a practical instrument useful for both business analysts and researchers, allowing them to detect, for instance, the determinants of high quality innovations.

Hence, this work constitutes a contribution to the analysis of innovation strategies by posing the following research question: *are patent data useful to provide a complete overview of innovation processes carried out by companies?*

1.3 Aim of the research

The research presented in this thesis proposes an integrated framework for defining innovation strategies carried out by companies through the analysis of information disclosed in patent data.

The proposed framework is based on the combination of five dimensions:

- technological strategy carried out for each knowledge domain involved in the innovation process;
- technological specialization within the technological field;
- openness of the innovation process;
- type of innovation resulting from R&D efforts;
- quality of innovation output.

Hence, the research objectives of my PhD thesis are the following:

O1: Literature review

To review the state-of-the-art of innovation research on each dimension under investigation, identifying the operationalisations mostly acknowledged by scientific literature.

O2: Patent-based framework for innovation strategy

To design a methodology that investigates each dimension and provides a more complete overview of innovation processes, also detecting the relationships among such dimensions.

O3: Validation

To test the methodology with an empirical study on top R&D spending companies belonging to bio-pharmaceutical and technology hardware & equipment industries, in order to validate both the framework applicability and its explicative power and usefulness.

1.4 Structure of the thesis

In the following section, I give a summary of the contents of my thesis, which is structured in five main parts:

Part I: Literature review. This part discusses the literature contributions about technological strategies, technological specialisation, openness of innovation processes, type of innovation and quality of innovation output. As such, the section attempts to summarize the state-of-the-art of innovation research.

Part II: Framework definition. In this part, the thesis will assume its experimental characteristic. Indeed, the patent-based framework for defining innovation strategies will be presented. Starting from the analysis of the operationalisations suggested by scientific literature, patent-based metrics are defined to summarise each dimension under investigation.

Part III: Framework application. This methodology will be validated through an empirical study on the sample under investigation. This part will be also dedicated to the evaluation of the results, reporting the statistical analyses. In addition, some challenges in implementing the metrics proposed will be underlined.

Part IV: Conclusions. This chapter includes the conclusions of the research: the major contributions, after both a theoretical and a practical perspective, will be provided. Limitations, future research lines and challenges will be also delineated.

Part V: Supporting material. This section comprises the appendices A and B, useful to deepen methodology and results. They are attached in the end of the thesis after references.

2. LITERATURE REVIEW

In this chapter, I aim at providing a theoretical overview on the issues addressed in the thesis, discussing literature contributions about knowledge management, innovation management and open innovation. Since the suggested framework consists of five dimension under investigation, I dedicated a specific literature review to each one.

2.1 Technological strategy

The technology of a firm is the result of its accumulated experience in design, production, problem-solving and trouble-shooting activities. Companies progressively accumulate their technological knowledge; therefore, firm's existing stock of knowledge is history dependent and affects its future technological development (Tsang, 1997). Technological innovation is featured by searching activities of optimal alternatives addressed to identify and solve technical problems (Nelson and Winter, 1982). Through such activities, companies may improve their current technological capability or develop new capabilities. Technological innovations are based on the recombination and the integration of capabilities belonging to different knowledge domains, therefore such processes depend on the experience accumulated by the company. Actually, companies operate using a wide range of knowledge domains and differ in their technological diversity (Pavitt *et al.*, 1989). Furthermore, in industries characterised by intense R&D activities, the competencies required to manufacture a product include multiple knowledge domains (Figure 1).

Therefore, companies pursue different innovation approaches that lead to different innovation performances, depending on the specific technological strategy adopted to achieve a competitive advantage in the industry in which each firm is involved. This process is not static, with companies expanding their breadth of knowledge over time (Pavitt *et al.*, 1989; Chang, 1996; Miller, 2004): knowledge does not have a rigid nature, but it can be transformed, accumulated, stored and transferred (Lo Storto, 2006).

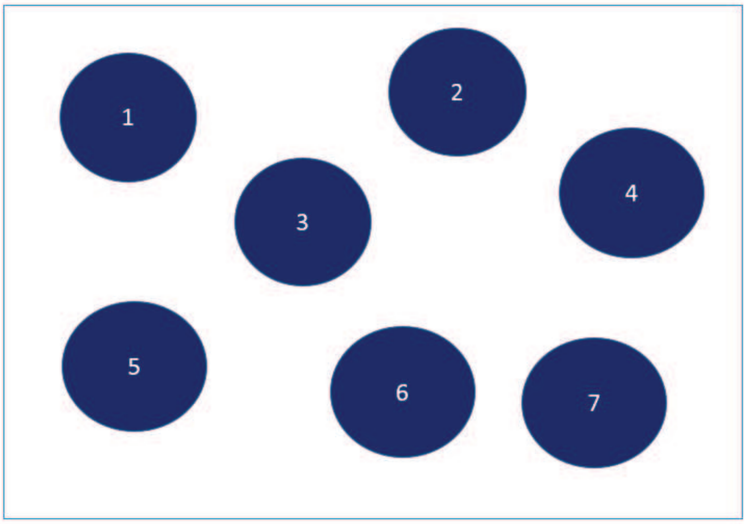

Figure 1. Stock of knowledge as the set of technological fields investigated by the company.

The dynamic evolution of capabilities is influenced by the exploitation vs. exploration strategy adopted by a company. March (1991) made an explicit distinction between exploration and exploitation; the former refers to the creation of new capabilities by means of activities such as fundamental research, experimentation, and search. The latter concerns the leveraging of existing capabilities by means of activities such as standardization, upscaling, and refinement. Figure 2 shows the effects of exploitation and exploration strategies on the stock of knowledge. Among the knowledge fields owned by a firm, in year t 1,3,5 and 6 are exploited, while 2,4 and 7 are not involved in new inventions, but they will still be part of the stock of knowledge because of the experience accumulated by the company, unless they will be abandoned in the following years. The orange bubble represents a new technological domain that is explored by the firm in year t. Therefore, exploration strategies extend the number of fields owned by the company.

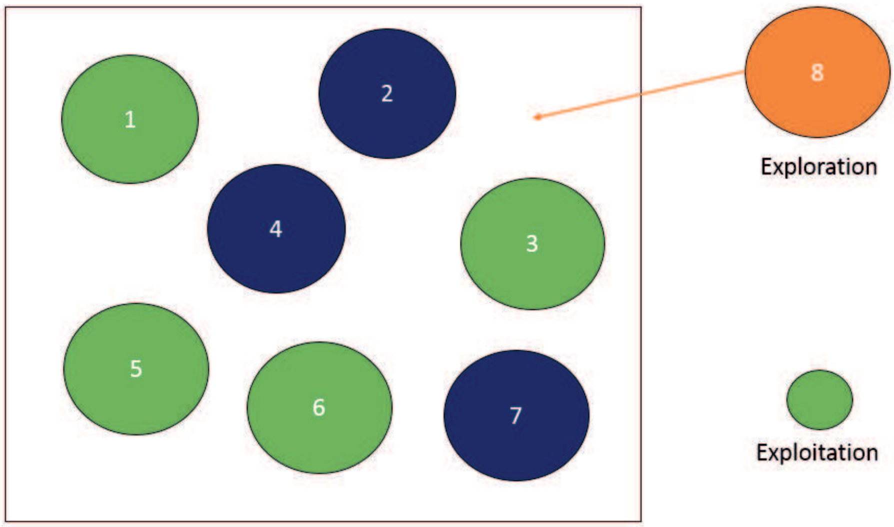

Figure 2. Effects of exploitation and exploration on the stock of knowledge.

Specifically, exploitation strategies are associated with experiential refinement, selection, reuse of existing routines, upscaling, standardization and recombination and are aimed at strengthening basic knowledge already owned by a company, and increasing the degree of novelty with a limited risk, within the boundaries of the present concepts and architectures (Simon, 1991). Exploitative activities improve the effectiveness and the efficiency of existing capabilities, require the creation of economies of scale and lead to short-term effects (Belderbos *et al.*, 2010). Exploitation, more often than not, generates incremental knowledge with moderate but certain and immediate returns (Schulz, 2001). Therefore, exploitative strategies are based on the local search and build on the existing technological trajectory, aiming at improving existing product-market domains (He and Wong, 2004). On the other hand, the exploration of new possibilities and ideas is based on distant search and associated with experimentation, play, risk taking, in order to both create new capabilities (Belderbos *et al.*, 2010) and produce new knowledge (Miner *et al.*, 2001). Such novel body of knowledge will serve as the seed for future technological development (Miller *et al.*, 2007) with companies involved in shifting to a different technological trajectory (Benner and Tushman, 2002) and aiming at entering new product-market domains (He and Wong, 2004) in order to achieve a long-term growth. Moreover, to execute distant search, a firm must identify distant knowledge domains from outside its boundaries and transfer them inside (Miller *et al.*, 2007). For instance, firms' members may attend conferences, browse patents, read trade journals or reverse-engineer competing products. Otherwise, companies may hire new skilled personnel in order to acquire their competencies in specific technological domains. In general, the capability to assimilate, assess and use new knowledge depends on firm's absorptive capacity (Cohen and Levinthal, 1990).

Even though exploration and exploitation compete for scarce organizational resources (March, 1991; 1996; 2006) and are self-reinforcing, adopting only one strategy may lead to a trap. Exploitation leads to early success, but often creates a success trap, with existing core capabilities turning into core rigidities, reducing the ability of the firm to adequately respond to technological changes and compromising the long-term survival (Leonard-Barton, 1992; Christensen and Overdorf, 2000). Differently, exploration often leads to failure and requires high commitment and investments, thus relying solely on it negatively affects firm's financial performance (i.e. failure trap). Therefore, firms benefit from a balanced mix of exploration and exploitation strategies and the combination of both strategies improves survival chances, growth and financial performances (O'Reilly and Tushman, 2004).

Table 1 summarizes the contribution provided by innovation literature about the definition of the different features regarding each technological strategy (Cohen and Levinthal, 1990; Simon, 1991; Miner *et al.*, 2001; Schulz, 2001; Benner and Tushman, 2002; He and Wong, 2004; Miller *et al.*, 2007; Belderbos *et al.*, 2010).

Key features	Exploitation	Exploration
Knowledge domains	Existing	New
Aims	Improve effectiveness and efficiency of current knowledge; Increase competencies and experience with the technology; Create economies of scale on the existing technological trajectory	Create a novel body of knowledge; Find new perspectives and more effective solutions to a given problem; Shift to a different technological trajectory
Activities	Standardization; Upscaling; Recombination; Reuse of existing routines; Refinement; Local search	Fundamental research; Experimentation; Play; Risk taking; Distant search
Product-market domains	Existing	New
Effects	Short-term growth; Specialization	Long-term growth; Heterogeneity in firm's problem-solving arsenal
Risks	Familiarity trap; Success trap; Rigidity trap; Learning trap; Competency trap	Information overload; Failure trap

Table 1. Exploitation vs. exploration strategy.

Gupta *et al.* (2006) suggested that, within a single technological domain, exploration and exploitation are mutually exclusive, whilst across different areas they are orthogonal, thus high levels of exploration or exploitation in a specific domain may coexist with high levels of exploration or exploitation on other ones.

2.2 Technological specialization

Although scientific literature takes into account the frequency with which companies operate within a specific technological field, no contribution was uncovered about the weight of such field on the overall innovation strategy carried out by firms. As a matter of fact, not all the domains are equally relevant for the company: only some knowledge areas are strongly stressed and mostly contribute to the development of the core technology of current business activities. Therefore, a different technological specialization may be detected in each technological field. Overall, companies concentrating their R&D efforts on few relevant fields are carrying out a focalization strategy. Conversely, working on a larger body of knowledge is related to differentiation strategies.

In literature concerning innovation management, few contribution take into account the trade-off between specialization and differentiation, therefore the topic is under-researched. A similar concept is expressed by the so-called technological familiarity, i.e. a component is familiar to the firm when it has been recently and frequently used (Fleming, 2001; Arts and Veugelers, 2012).

2.3 Openness of the innovation process

Regarding the organizational dimension of R&D, companies can carry out the entire innovative effort within their boundaries, pursuing a closed innovation (CI) strategy, or open up their R&D processes, adopting an open innovation (OI) approach.

2.3.1 Closed vs. open innovation

In a closed system, new products and services are developed in-house and exploited by the company to enter the market first and win. After an open perspective, innovations are the result of collaboration efforts with third parties. By carrying on a CI strategy, capabilities and technologies are developed within the boundaries of the innovating firm and improved in order to reach the market and generate revenues. On the contrary, by implementing the OI strategy the boundaries of the innovation funnel become permeable (Chesbrough, 2003) with R&D projects a) jointly developed with other parties, b) developed by third parties before entering into the funnel or c) started by the company but leaving the funnel and further developed by third parties. The OI paradigm is conceived on the idea that companies are unable to hold in-house all the competencies they require, thus forcing them to open up their R&D processes. In the last decade, inter-firm R&D collaborations, strategic technologies alliances, joint development with universities and research groups, complex innovation networks and joint venturing investments have been incorporated into companies' technological strategies, since they give access to different knowledge bases and new resources.

Even though, in the last decade, inter-firm R&D alliances have become crucial for many companies belonging to industries featured by intense R&D activities, firms still prefer to develop internally their core products, without collaborating or outsourcing. Actually, they can be produced better, faster, and more efficiently internally and alone than in collaboration with other companies. Firms protect their invention with intellectual property rights and exploit the results of their R&D efforts in order to gain competitive advantage. Indeed, companies are rebuilding their internal R&D organization: already twenty years ago, Archibugi and

Pianta (1996) found that the importance of large in-house laboratories was declining and the international dimension of the innovation process was growing. The increasing geographic distribution of R&D activity is still an important phenomenon of globalization (Lahiri, 2010): firms decentralize their R&D activities and even more involve their subsidiaries in the technological development.

Even though centralized R&D can generate technologies of greater impact (Argyres and Silverman, 2004), firms achieve exploitation through specialization, dividing themselves into various units to focus efforts on specific products and geographic markets (Miller et al, 2007). The higher the complementarity of assets needed to bring products to market, the greater the divisionalisation of a company. Firms may geographically distribute their R&D activities in order to share and allocate different technological domains (Nayyar and Kazanjian, 1993; Nerkar and Roberts, 2004). Further, divisionalisation might result from merger and acquisition activities: in order to avoid the loss of the capabilities of the new subsidiary, it is usually best to allow it to remain intact.

In addition to their local inputs, firms may source knowledge from distant units (Venaik et. al., 2005). Such inputs comprise new knowledge both created in the distant unit and sourced externally by the distant unit. According to Lahiri (2010), with increasing geographic distribution of R&D activities, two issues may be defined: a) search costs increase, creating diseconomies of scale and b) transferring new knowledge from one unit to another becomes challenging.

2.3.2 Equity vs. non-equity alliances

Among the different OI practices, a distinction can be made between equity alliances and non-equity alliances. Within equity alliances (i.e. joint ventures) companies agree to share capital, technology, human resources, risks and rewards and establish a new entity under shared control. From a knowledge perspective, equity alliances provide the highest level of partner interaction and are considered as the most effective means of knowledge transfer (Anand

and Khanna, 2000). The joint creation of new knowledge requires high levels of resource commitment (e.g. capital, employees, time), equal motivation from both firms, and appropriate control mechanisms (Kogut, 1988). Conversely, non-equity alliances (i.e. R&D collaborations) are characterized by lower resource commitment and give access to new knowledge bases, ideas and possibilities through the interaction with partners (Granovetter, 1973). Collaborations may differ in frequency and duration of the relationship and number of partners. Therefore, in an OI system companies establish a complex inter-firm network of relationships with other organizations, in which each one teams up to generate new products and technologies (Dittrich and Duysters, 2007).

Many scholars studied the relationship between alliances and exploitative vs. explorative strategies. As to exploitation, intense collaborations with partners are required in order to achieve recurrent and trustful relationships (Krackhardt, 1992) and the creation of economies of scale. Companies pursuing an exploitation strategy will search for firms with similar technological capabilities: the collaboration needs time to build up and generates long-term benefits. Therefore, joint venturing strategies are mainly adopted in exploitative activities (Koza and Lewin, 1998), since companies need to establish strong ties with their partners and strong legal agreements. As of explorative strategies, they are pursued through alliances with partners with different capabilities, which give access to a different knowledge base. When exploring new technologies, firms need a more flexible form of alliance, since the result of the partnership is typified by more uncertainty and they need to have the possibility of abandoning the alliance at any given moment (Duysters and De Man, 2003). Furthermore, explorative activities require a continuous scanning of new technological opportunities. Since these opportunities often arise outside existing partners, partner turnover will be high (Dittrich and Duysters, 2007). For such reasons, non-equity alliances are strongly preferred in exploration strategies.

Given that collaborative R&D activities are characterised by a larger field of application, the adoption of non-equity alliances is expected also in exploitation strategies. Actually, many scholars discovered that companies jointly develop new products in order to share the costs of exploiting a certain form of technology (Nakamura *et al.*, 1996) and share risks and costs of innovation under growing technological complexity (Hung and Tang, 2008). In addition, in industries featured by high market fragmentation companies with similar core business activities collaborate in standardization consortia, setting the standard for a particular technology (David and Steinmueller, 1995; Schmidt and Werle, 1998; Egyedi, 1999; Hawkins, 1999).

2.3.3 R&D collaboration

Companies can build their stock of knowledge not only through internal efforts, but also opening up their R&D processes. The OI paradigm (Chesbrough, 2003) is conceived on the idea that when the innovation funnel boundaries become permeable, firms are able to explore and exploit technologies by sharing their innovative processes with third parties. Therefore, OI describes a new opportunity for a firm's strategy to profit from innovation, suggesting that companies should use inflows and outflows of knowledge to foster internal R&D, and to expand markets for external use of innovation, respectively (Chesbrough, 2006). OI has become one of the most addressed research topics in innovation studies, and many scholars investigated the impact of such strategy on companies' stock of knowledge. In particular, when internalization strategies are pursued, external know-how can be acquired from third parties or shared within collaborative development processes (Schroll and Mild, 2011). Different business models can emerge from OI adoption, ranging from R&D collaborations to the incorporation of knowledge-intensive firms (Michelino *et al.*, 2015a).

This work focuses on the collaborative dimension of OI, by investigating the impact of R&D collaboration activities on firms' stock of knowledge and innovation performances. Actually, for firms operating in dynamic environments - featured by rapid development and increasing knowledge complexity - it is very difficult to contain and capitalize on all relevant knowledge. Therefore, companies specialize and employ R&D collaborations to complement their knowledge (Perez *et al.*, 2013). This leads to the creation of complex networks of relationships with customers, suppliers and other public and private organizations, in which companies team up to generate new products and technologies (Dittrich and Duysters, 2007). Indeed, firms may potentially learn from such interactions, which open up opportunities for joint value creation and innovation (Prahalad and Ramaswamy, 2004). Therefore, the joint development can be a source of competitive advantage, since it gives access to external sources and information (Belderbos *et al.*, 2004). For these reasons, R&D collaborations and strategic technological alliances are increasingly part of companies' innovation model (Archibugi and Pianta, 1996). However, many difficulties in managing the complexity of such relationships occur; therefore, the share of alliance failures is significant (Patzelt and Shepherd, 2008). Thus, a key determinant for a successful collaboration is the absorptive capacity, which is "the ability of a firm to recognize the value of new, external information, assimilate, and apply it to commercial ends" (Cohen and Levinthal, 1990).

In summary, the reasons why companies enter into R&D collaboration are various:

- it can help firms to reduce uncertainty in terms of costs and risks (Tyler and Steensma, 1995; Das and Teng, 2000), share the costs of exploiting a certain form of technology embodied into new products (Nakamura, 2003), share risks and costs of innovation under growing technological complexity (Hung and Tang, 2008);
- it is required to shorten innovation cycles (Pisano, 1990) and enter the market first and win;
- it is necessary to define regulations and industry standards more effectively (Benfratello and Sembenelli, 2002);
- it gives access to new knowledge bases, ideas and possibilities through the interaction with partners (Granovetter, 1973);
- it is source of complementary expertise for companies operating in industries typified by technological complexity, where no single firm possesses all the knowledge, skills and techniques required (Powell *et al.*, 1996; Rausser, 1999).

Many scholars studied the determinants of partners' selection and defined different factors that affect the collaboration performance, such as: complementary skills and capabilities, cooperative culture, compatible goals, technological expertise, marketing system and knowledge, competitive strength, production efficiency, positive prior experience, labour negotiation experience, intangible assets, prior ties with universities (Brouthers *et al.*, 1995; Rausser, 1999; Nielsen, 2003; Wu *et al.*, 2009).

2.3.4 Industrial vs. scientific partnerships

Firms may collaborate with both industrial and scientific entities, such as other firms, universities and research centres. The most relevant advantages in collaborating with industrial partners are related to the access to complementary assets critical to successful development and commercialization, such as (Arora and Gambardella, 1990; Pisano, 1990; Teece, 1992):

- market access;
- marketing and distribution channels;
- production facilities;
- expertise in managing R&D and development processes;
- financial capital to support the focal firm;
- experience in evaluating payoffs far in the future;
- experience on how to operate and grow a firm in the same industry;
- strategic and operational know-how.

Furthermore, such R&D alliances allow companies to:

- possess stable exchange relationships (Stinchcombe, 1965);
- acquire innovative capabilities (Shan *et al.*, 1994);
- achieve external endorsement of their operations (Baum and Oliver, 1991);
- employ the perceived quality and reliability of their products and services among potential customers, suppliers, employees, collaborators and investors (Stuart, 1998).

On the other hand, firms provide their capabilities in new technologies and, then, their specialization, useful for their partners. For example, many biotech companies completely rely on OI as the unique source of revenues, by operating as innovation sellers with larger pharmaceutical firms (Michelino et al, 2015b).

Regarding relationships with universities, research institutes, government labs and hospitals, various advantages in undertaking a collaboration can be defined:

- the partners are a source of up-to-date information and knowledge, which is too tacit to be transferred through licensing or acquisition (Liebeskind *et al.*, 1996);
- collaboration with universities gives access to international knowledge networks (Okubo and Sjöberg, 2000) and, consequently, to international markets;
- universities and scientific partners can be involved in developing prototypes and handling patents and licenses (Cyert and Goodman, 1997);
- by collaborating with public entities, companies may access to public resources and funds (Bayona Saez *et al.*, 2002). Therefore, the collaboration is stimulated by public programs which promote research and partnerships between public and private search;
- R&D collaborations provide interaction opportunities, which generate new concepts, business ideas, emerging knowledge and technological know-how that firms can translate into new products (Powell *et al.*, 1996);
- R&D partners can give an answer to the demand for both basic knowledge and pre-competitive research (Arora and Gambardella, 1994) and more specific knowledge, focusing on problem solving and product design and development, i.e. applied research (Bayona Saez *et al.*, 2002);
- partnerships with universities and research centres allow companies to keep up-to-date in industrial standards and to access government information useful to find out what other firms in the sector are doing (Sakakibara, 1997).

R&D partnerships with universities and other scientific organizations are usually featured by long-term collaborations because of the basic and complex nature of the joint research, which requires larger learning processes (Hall *et al.*, 2000a). In these partnerships, companies need to be particularly able to absorb the scientific knowledge transferred by partners and have a strong internal capacity for R&D (Bayona Saez *et al.*, 2002). In many cases, the results obtained are not directly exploitable for business applications. Another relevant feature characterizing such collaboration is the geographical proximity, which significantly improves the efficiency of the alliances (Baptista and Mendonça, 2000). Actually, it simplifies and reduces the R&D efforts, which are affected by differences in aims, management styles and culture.

2.4 Type of innovation

According to Henderson and Clark (1990), innovation can be defined from the analysis of two features: changes on core concepts and impact on the linkages between core concepts and components (Figure 3).

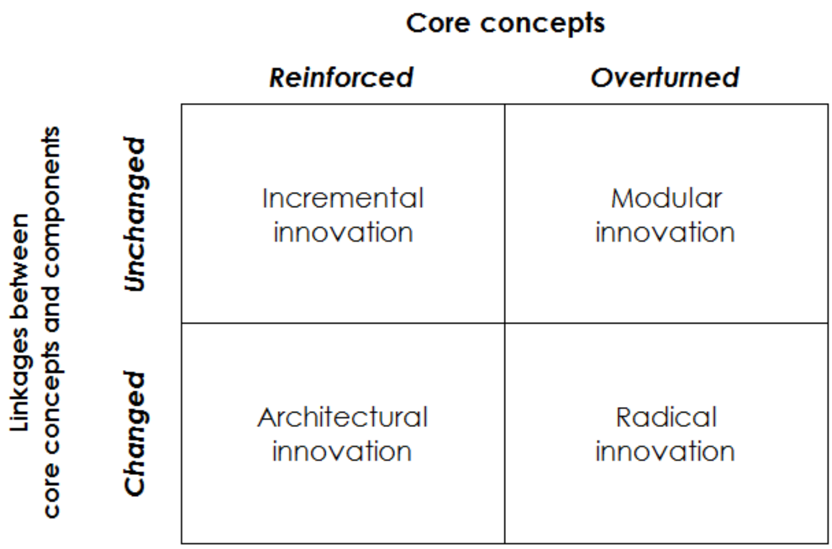

Figure 3. Types of innovation (Henderson and Clark, 1990).

The former refers to the novelty level of the innovation and has been widely explored in literature. Ahuja and Lampert (2001) termed technologies without technological antecedents as pioneering. Such technologies focus on completely *de novo* solutions. Actually, inventions without technological antecedents can be considered pioneering (Trajtenberg *et al.*, 1992; 1997; Ahuja and Lampert, 2001; Rosenkopf and Nerkar, 2001; Shane, 2001; Kaplan and Vakili, 2012), while the existence of technological antecedents is a proxy of innovations based on the reinforcement of core concepts (Jaffe *et al.*, 1993; Hall *et al.*, 2001).

Regarding the impact of the linkages between components, innovation can be achieved through recombining already established elements (Fleming, 2001) or by introducing an established element into a new setting (Hargadon and Sutton, 1997). Henderson and Clark (1990) argue that the mere rearrangement of previously used components can itself cause destabilizing industrial change. Therefore, new technologies can derive from the combination of elements and settings not previously observed (Dahlin and Behrens, 2005). Thus, companies with most relevant skills and technology from prior related R&D can be capable of recombining these resources and capabilities in a novel way and originate innovation, i.e. recombining existing but disconnected technology components can foster innovation (Arts, 2012). The major technological shifts frequently emerge from the fertilization between different pre-existing but disconnected technology subfields or components (Basalla, 1988; Hargadon, 2003). By recombining different technology components, inventors search for those solutions with the highest fitness value (Fleming and Sorenson, 2001).

From the matrix proposed by Henderson and Clark (1990) four types of innovation can be defined. When core concept are reinforced and linkages between core concepts and components are unchanged, the output of the R&D effort is an incremental innovation (Figure 4). Incremental innovations occurs also when only part of the components of an architecture are technologically improved, without changes in the architecture. Incremental innovation refines and extends an established design. Improvements occur in individual components, but the underlying core design concepts, and the links between them, remain the same (Henderson and Clark, 1990). These incremental technological improvements enhance and extend the underlying technology and thus reinforce an established technical order (Tushman and Anderson, 1986).

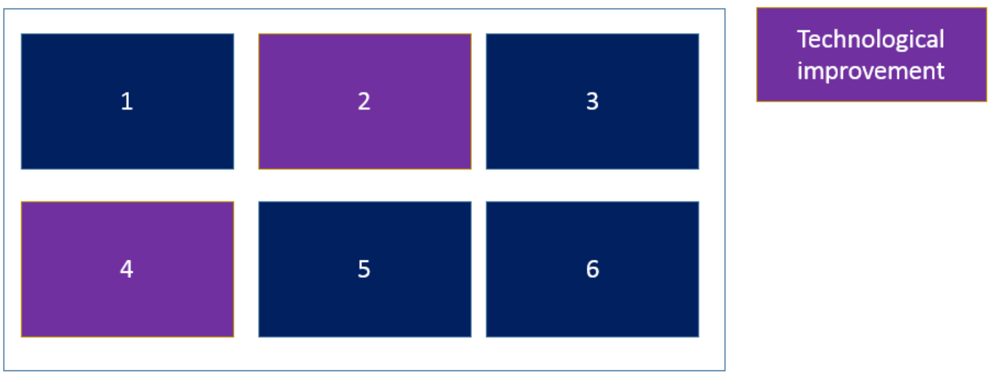

Figure 4. Incremental innovation.

On the other hand, when the components/modules are overturned and a new technology is employed within the architecture, a modular innovation is obtained (Figure 5). The innovation relates to the introduction of new component technologies in specific parts of the product, without effects on the overall architecture. Thus, this implies that the innovation primarily affects isolated parts of the product system and that the technological novelty is concentrated to these parts.

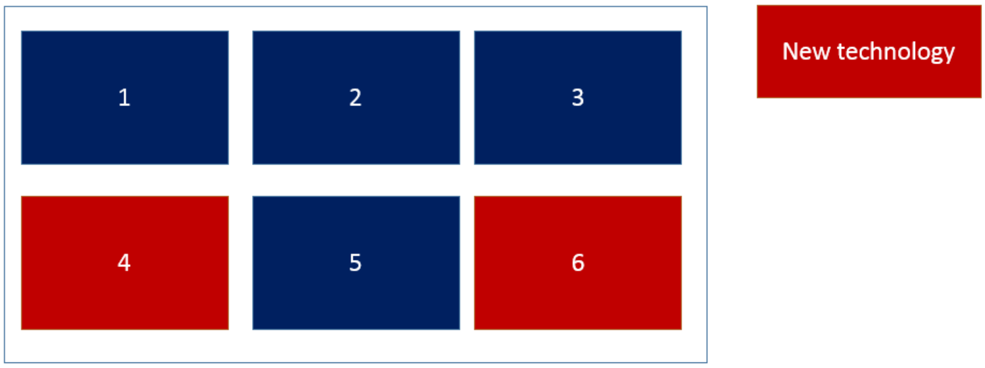

Figure 5. Modular Innovation.

Regarding architectural innovation (Figure 6), it consists in the creation of new combinations of components (i.e. new settings), without using new technologies or core concepts; at most they are reinforced (even partially). Architectural knowledge concerns the ways in which the components are integrated and linked together into a coherent whole. Therefore, architectural innovation destroys the usefulness of a firm's architectural knowledge but preserves the usefulness of its knowledge about the product's components. The essence of an architectural innovation is the reconfiguration of an established system to link together existing components in a new way. This does not mean that the components themselves are untouched by architectural innovation. Architectural innovation is often triggered by improvements in a component (e.g. size or some other subsidiary parameter of its design) that creates new interactions and new linkages with other components in the established product. The important point is that the core design concept behind each component and the associated scientific and engineering knowledge remain the same.

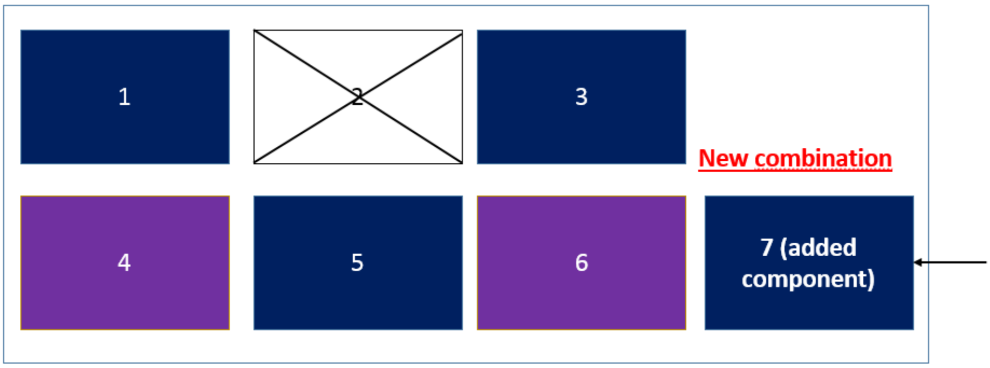

Figure 6. Architectural innovation.

Finally, radical innovation is achieved when new technologies are combined with other components, in order to create a new combination, i.e. an architecture that strongly departs from past settings (Figure 7). Therefore, radical innovation is based on a different set of engineering and scientific principles and often opens up completely new markets and potential applications. Radical innovations establish new dominant designs and, hence, a new set of core design concepts embodied in components that are linked together in a new architecture (Henderson and Clark, 1990). They are fundamental changes that represent revolutionary change in technology and represent clear departures from existing practice.

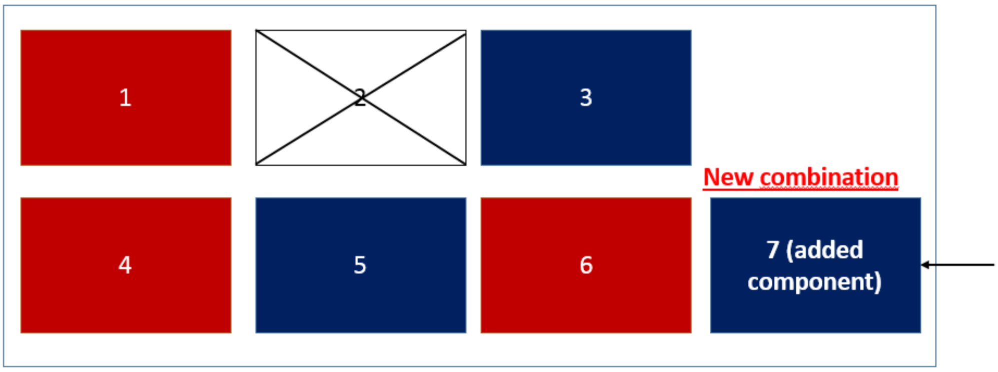

Figure 7. Radical innovation.

2.5 Quality of innovation output

The quality of innovation output can be regarded as the impact of the new technology. At a very basic level, a distinction can made between inventions that impact from a technological perspective vs. inventions that are shocks from a user or market perspective (Ahuja and Lampert, 2001).

From the technological point of view, inventions can be related to the concept of breakthrough innovation. Breakthroughs offer high improvements in the price vs. performance ratio over existing technology (Tushman and Anderson, 1986) and are generally considered as competence-destroying, since they are based on science and engineering principles rendering obsolete knowledge, capabilities and technology previously accumulated. Such inventions serve as the basis for many subsequent technological development (Trajtenberg, 1990a; 1990b) and can be regarded as breakthrough because they have demonstrated their utility on the path of the technological progress (Fleming, 1998).

Nevertheless, achieving such breakthroughs may not be wholly consistent with achieving economic impact, since a technology embodied in an invention requires several other factors, such as the reputation of the firm, the distribution of the idea in the network, the presence of complementary technologies (Kaplan and Vakili, 2012). Similarly, not all marketable innovations are necessarily technological breakthroughs: inventions that make an original idea more understandable and usable or that distribute it strategically in a network may also have success.

3. METHODOLOGICAL FRAMEWORK

Starting from a literature review and focusing on the operationalization on patent data that are already acknowledged and adopted by scholars, I designed a patent-based framework with the aim of investigating the five dimensions of innovation previously defined and defining the relationship among them.

Indeed, patent data are the only formally and publicly verified outputs of inventive activities and are widely accepted as a measure of innovation. As suggested by Griliches (1990), data provided by patents contain information about the whole population of innovating firms, are standardized, stored for a long period and continuously updated. Patent statistics provide very specific and detailed information for evaluating inventive activities (Acs and Audretsch, 1989; Chakrabarti, 1991; Grupp, 1992). Furthermore, they are objective, since they have been processed and validated by patent examiners (Belderbos *et al.*, 2010). In addition, they capture the proprietary and competitive dimension of technological change (Archibugi and Pianta, 1996). Since obtaining patent protection is costly and time-consuming, only inventions that are expected to provide benefits that outweigh cost are applied. For all these reasons, patents have been treated as the most important output indicators of innovative activities and patent data have become the focus of many tools and techniques to measure innovation (Ma and Lee, 2008). Unlike other innovation metrics (e.g. R&D expenditures, number of R&D personnel) which regard the input of R&D activities, patent data focus on outputs of the inventive process, and more specifically of those inventions that are expected to have a commercial impact. They provide a valuable information about the effects of technological innovation and can be disaggregated to specific technological domains (Johnstone *et al.*, 2012). For the analysis of specific development activities carried out by companies at the technological domain level, scholars refer to patent data for studying the relationship between innovation and knowledge (Almeida and Kogut, 1999; Abraham and Moitra, 2001; Ahuja and Katila, 2001; Ahuja and Lampert, 2001). When they need to analyse knowledge domains, they focus on

International Patent Classification (IPC) codes, which identify the belonging technological fields of an innovation. Actually, all patents are categorised into at least one IPC: such technological index operates like a keyword system (Graff, 2003). IPC codes are widely employed to investigate technological innovation strategies implemented by innovative firms: e.g., Sakata *et al.* (2009) studied IPC combinations in order to define the innovation position of Japanese companies, while Suzuki and Kodama (2004) described technological trajectories and technological diversification strategies by examining patent classification codes.

In this thesis, the analysis is performed at the knowledge domain level, evaluating the different innovation strategies adopted within each technological field involved in firms' R&D efforts. Furthermore, the behaviours detected are taken into account in order to define the overall innovation strategy pursued by each company. In this thesis, I tested the framework on a sample of firms belonging to the bio-pharmaceutical (BP) and technology hardware & equipment (THE) industries.

3.1 Priority patents and patent data

The work examines the claimed priorities from the population of patent applications recorded in PATSTAT database, since these are considered the high-value applications (Harhoff *et al.*, 2003; Johnstone *et al.*, 2012). A priority filing is the first patent application filed to protect an invention. It is generally filed in the patent office of the inventor's country of residence, although it may be applied elsewhere. Such filing is followed by a series of subsequent filings and together they form a patent family (Figure 8). Since such subsequent applications are the consequence of the first effort carried by a firm during the first filing, I refer only to claimed priorities. Furthermore, by considering only priority filings I can avoid to involve within the analysis multiple applications related to the same invention. Moreover, the priority date is the one closest to the period in which R&D activities were completed, thus supporting me to better define the reference year for the invention.

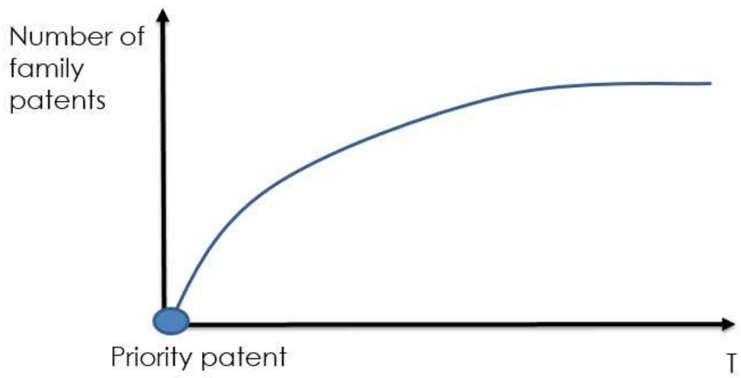

Figure 8. Example of family size growth over time.

Only patents filed at the EPO and USPTO offices were taken into account. I employed a broad set of patent information to model the respective variables:

- Cooperative Patent Classification (CPC) codes;
- list of companies disclosed in the assignee field;
- backward and forward citations;
- family size;
- renewal fees.

In what follows, for each dimension under investigation I summarise the patent-based metrics mostly acknowledge by scholars and, starting from their operationalisations, I present how to implement the analysis within the following dimensions:

- technological strategy;
- technological specialization;
- openness of the innovation process;
- type of innovation;
- quality of innovation output.

3.2 Technological strategy

According to Belderbos *et al.* (2010), technological domains can be analysed in order to evaluate companies' exploitation vs. exploration strategies: a patent is considered as explorative if it is situated in a technological domain in which firm lacks of prior familiarity. Therefore, explorative innovation activities develop ideas situated in knowledge fields where the firm has not patented in the past five years, whilst exploitative ones refer to technologies developed in knowledge areas where the firm has patented technology in the previous five years. This assumption is in line with the idea that knowledge evolves rapidly and companies lose most of their technical experience if they abandon a technological field for five years (Argote, 1999; Ahuja and Lampert, 2001; Fleming, 2001; Hall *et al.*, 2005; Leten et *al.*, 2007), with competencies previously accumulated resulting obsolete and forcing them to re-explore such technological domain. Furthermore, when companies start to explore a new knowledge area, it remains relatively new until they accumulate experience in the search activity within it: such process requires time and resources; therefore, Belderbos *et al.* (2010) suggest that a technological field keeps its explorative status for a period of three consecutive years.

Starting from this operationalization, I analysed classification codes disclosed in patent documents. For each analysed company I downloaded patent data from PATSTAT database, considering patents applied in the investigated time interval and detecting their classification codes. Even though scholars examine technological fields through IPCs, in this framework I refer to the Cooperative Patent Classification (CPC) system, a nomenclature developed by the European Patent Office (EPO) and the United States Patent and Trademark Office (USPTO) in order to allow inventors to retrieve relevant prior art efficiently. Actually, such system combines the best practices of the two offices and was built starting from IPC classification; therefore, it may be considered as its evolution, since it is more specific and detailed: while the IPC has about 70,000 entries, the CPC has more than 250,000 ones, making it much more precise. The standardization allows to analyse patent applications with both the EPO and the USPTO as a receiving office. Each CPC consists of a hierarchical symbol: the first letter defines the section, the two digits number denotes the class and the following letter identifies a sub-class. The sub-class is then followed by a one to three digits main group number, an oblique stroke and a number representing the sub-group. Unlike the operationalization applied in literature, I decided to cut the code and consider only the information before the stroke, since I believe that the operationalization of knowledge fields requires more generalization. For instance, Table 2 displays the hierarchical composition of the CPCs "H04W88/08" (i.e. access point devices) and "H04W88/12" (i.e. access point controller devices) with the interpretation about the meaning to be assigned for research purposes.

Level	Symbol	Hierarchy	Classification	Meaning
1	H	Section	Electricity	Technological base
2	H04	Class	Electrical communication systems	Technological sector
3	H04W	Sub-class	Wireless communications networks	Technological segment
4	*H04W88*	Main group	*Devices specially adapted for wireless communication networks*	*Knowledge domain (Technological field)*
5	H04W88/08 H04W88/12	Sub-group Sub-group	Access point devices Access point controller devices	Products or components

Table 2. Example of CPC hierarchical composition.

By considering the entire code (level 5), I may study innovation at the component level, or rather at the maximum level of disaggregation. Since I aim at analysing innovative behaviours at the knowledge domain level, I require a higher level of aggregation and decide to consider the code at level 4 (i.e. until the main group). For example, both CPCs shown in Table 2 belong to the same technological field (i.e. devices specially adapted for wireless communication networks): I hypothesize that different products or components may be developed within the same knowledge domain since competencies required in the innovative process are almost the same for both. Similarly, an excessive level of aggregation does not allow to correctly identify the various capabilities that a company owns. For each firm belonging to the sample, I detected from PATSTAT database the distinct CPCs disclosed in its patent applications. Each technological field is then labelled as exploitative or explorative. I started from the operationalization suggested by Belderbos *et al.* (2010):

- a knowledge domain is labelled as exploitative if the company filed patents in such technological field in the past five years, explorative otherwise;
- the technological field keeps its explorative status for a period of three consecutive years.

The five-year time span is based on the assumption that companies lose their previous experience if they abandon a specific technological domain, while the three-year one, used for evaluating the exploration, is necessary for companies to master a knowledge field before it is exploitable. Yet, such hypotheses do not seem to take into account the different features of the belonging industry of companies. For instance, in the BP sector the development of a new drug can take more than five years: the lack of patent applications in a specific technological domain in the previous five years does not imply the loss of knowledge, since an invention may be in the development phase. Hence, the experience interval should consider the higher time-to-prototype and, thus, should be increased. On the other side, in the THE industry the faster development pace and the shorter product life cycles force companies to continuously adapt their technical competencies, which may be considered obsolete in a time span lower than five years. Thus, in order to take into account industry-specific time spans, I adjusted the experience interval:

- by adding 2 years for companies belonging to the BP industry, resulting in a seven-year time span;
- by removing 2 years for firms belonging to the THE one, considering a three-year time span.

Consequently, also the exploration interval is influenced by industry-specific characteristics (e.g. product complexity and development pace), with BP companies requiring more time to make a technological field exploitable and THE ones forced to speed up the process of familiarization with a new knowledge domain. Therefore, the exploration time span is set at:

- 4 years for companies in the BP industry;
- 2 years for THE firms.

Such operationalization is in line with the different market, product and industrial structures in which companies compete (Table 3). Without accounting for the time span adjustment factors, a comparison between the two industries may lead to inaccurate results. For this reason, I adjusted the values recognised in the scientific literature, which I consider as mean values applicable to all industries, rather than building them ex-novo.

Industry characteristic	**BP industry**	**THE industry**
Product development time	About 10-12 years[1]	About 1-3 years
Research activity	Basic research	Application science and engineering
Regulations	Government regulations	Industry standards and customer expectations
Product	Integral nature	Modularity of IT design, component-based products
Uncertainty of R&D process	High	Medium-low
Products and Intellectual Property Rights	Product covered by a small number of patents	Many patents to assemble intellectual property rights for a single product
Patenting strategy	The company is the sole holder of a drug patent	The firm holds just a large enough percentage of the total relevant patents

Table 3. Market, product and industrial structure for BP and THE companies.

[1] Since bio-pharmaceutical companies have to apply the patent before a drug is subject to the evaluation of public health authorities, I assume that the invention is filed within 7 years from the beginning of the project, in line with the operationalization of the experience interval.

In order to label as exploitative or explorative each distinct CPC detected for every company, an examination of previous patent applications is required. Such study is limited by the experience interval defined for the belonging industry of the company. This range of time can be divided into two periods: the exploitation phase and the exploration one. Since the latter is previously defined and industry-specific, the former is fixed by difference. For instance, in the BP industry I analysed only patent applications from t-7 to t-1, having considered a seven-year time span for the of experience interval: the period of exploration is 4 years, then the time span for the exploitation is 3. I am supposing that the knowledge owned by the company before t-7 is no more useful and available in t, if afterwards it was not further accumulated and recombined, bringing to a new patent application. Since a CPC is labelled as exploitative if the company has already patented within the knowledge domain and has already trespassed the exploration phase, in the BP industry only technological fields for which at least a patent application is detected from t-7 to t-4 can be considered as exploitative in t. However, if no patent was applied in such time interval, the technological field is still in the exploration phase: even though a patent application is detected from t-3 to t-1, I assume that in t within the knowledge domain the exploration phase is not yet complete (Figure 9). Therefore, if no patent application reporting the specific CPC was found from t-7 to t-1, the CPC is new for the company, since the first patent has been applied in t and the knowledge domain is labelled as explorative.

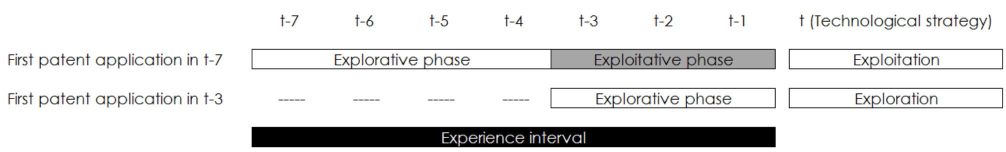

Figure 9. Labelling of knowledge domains from patents filed by BP companies.

As to the THE industry, I considered only patents from t-3 to t-1, and the CPC is labelled as exploitative if I find applications in t-3, explorative otherwise (Figure 10).

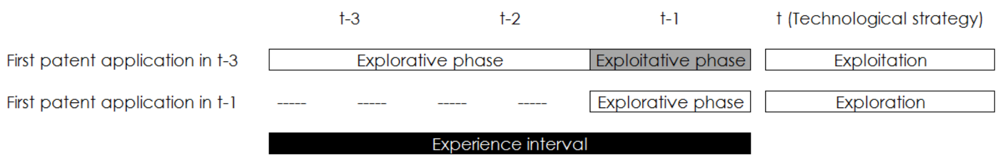

Figure 10. Labelling of knowledge domains from patents filed by THE companies.

Hence, I investigated the features of each knowledge domain involved in each invention and I summarize such features through shares defining two variables:

- EXPLOIT, as the ratio of CPCs in the exploitation phase declared within the patent application;
- EXPLOR, as the share of CPCs in exploration phase recorded in the document.

For instance, for a patent declaring four CPCs, of which three are labelled as exploitative, EXPLOIT is 75% and, consequently, EXPLOR is 25%.

3.3 Technological specialization

After having labelled a CPC as exploitative or explorative, a second label is assigned: core vs. non-core. Since not all the domains are equally relevant for the company, only some knowledge areas are strongly stressed and mostly contribute to the development of the core technology of current business activities. I aim at distinguishing between inventions in core technological fields and those in non-core ones and evaluating if the different relevance significantly affects the choice of management and organization of innovation strategies. Also technological specialization is analysed using the 4-string CPC code. In particular, each CPC is defined as core if it is declared in at least 10% of the patents filed in the experience interval, non-core otherwise[2]. For instance, in the BP industry the relevance is estimated dividing the number of patents declaring the analysed CPC from t-7 to t-1 by the total amount of patent applications recorded from t-7 to t-1, considering only those reporting at least one CPC. This idea is based on the assumption that a technological field is core if its accumulation of knowledge in the experience interval generates a large number of patent applications. As in the case of technological strategy, a ratio of core and non-core technological fields is defined for each patent, named respectively CORE and NONCORE.

In literature, a similar concept is expressed by the so-called technological familiarity. A component is familiar to the firm when it has been recently and frequently used (Fleming, 2001; Arts and Veugelers, 2012). For each separate technological field disclosed in a patent document, scholars estimated an individual measure of component familiarity and the average component familiarity of the patent. Therefore, I deduce that familiarity can be seen as the absolute value from which I may derive a relative measure (i.e. technology specialization).

[2] The threshold of 10% is based on robustness tests. Indeed, by setting up the threshold to 15% for many companies no core knowledge domains were found, whilst reducing it to 5% the majority of technological fields is labelled as core, impeding a good distinction between core and non-core CPCs.

3.4 Openness of the innovation process

Regarding the management and organization of innovative activities, researchers focus on the assignee field disclosed in patent documents to define the openness of the process. When a firm develops in-house a new technology, only one applicant is recorded in the patent application. On the contrary, a co-assignment is detected when two or more companies are involved in the development and contribute to the final invention, sharing the ownership of the innovation. Thus, co-patents seem to be a relevant indicator for signalling the occurrence of OI strategies (Chesbrough, 2006) and the number of patents deriving from collaborative projects can be considered as a proxy of OI (Al-Ashaab *et al.*, 2011). Kim and Song (2007), using joint-patenting information, reported a growing OI adoption.

In order to delineate how companies manage and organize their R&D efforts, I refer to the assignee field disclosed in patent documents. Through the analysis of such a field, I am able to define which are the actors involved in the development of the invention.

The first step of the analysis regards the linkage between the companies of the sample and the PATSTAT applicant table. For each firm I searched in the assignee field both the name of the parent company and its subsidiaries, disclosed in the 2011 annual report, also taking into account the names of the units previously acquired or merged.

The second step of the analysis refers to the study of the companies disclosed in the applicant field. Since I left out any inventor from the framework, only firms are considered. I focus on the number of assignees recorded in patent documents in order to find information about the management and organization of the specific invention. Particularly:

- if only one company is found, the patent is internally developed;
- if two or more different organizations were found in the applicant field, I detect a joint patent among the analysed firm and third parties.

Regarding the partner typology, each partner is labelled as:

- industrial company, if the patent document records a private organization;
- scientific organization, when relationships involve universities, research institutes, government labs, hospitals and any other public authority.

The purpose is to understand if the different partner typology features specific firms and affects other characteristics under investigation. Therefore, three indicators can be defined for each patent:

- CLOSED, assuming value 1 if the patent has been internally developed, 0 otherwise;
- OPEN_SCI, a dummy variable signaling the occurrence of scientific organizations on within the focal patent;
- OPEN_IND, equals to 1 if at least an industrial partner has been detected for the focal patent.

3.5 Type of innovation

Consistently with Henderson and Clark (1990), two features are investigated in order to identify the type of innovation: the novelty level of an innovation and the impact on the linkages between components.

The former can be investigated through the analysis of backward citations. Actually, patents without backward citations to prior technical art can be considered pioneering (Trajtenberg *et al.*, 1992; 1997; Rosenkopf and Nerkar, 2001; Shane, 2001; Kaplan and Vakili, 2012), while the existence of backward citations is a proxy of innovations based on the reinforcement of core concepts (Jaffe *et al.*, 1993; Hall *et al.*, 2001). Therefore, the lack of prior art citations is a proxy of originality, creativity and novelty of a patent. On the contrary, patents disclosing backward citations can be considered as spillovers. Even though the applicant may deliberately avoid backward citations, he has the legal duty to disclose any citation to prior art. The decision regarding which patents to cite ultimately rests with the patent examiner, who is an expert in the area and able to identify relevant prior art that the applicants omitted. Furthermore, applicants may cite their previous inventions, signalling an internal spillover and the reinforcement of their own core concepts.

A significant contribution regarding such operationalization derives from Jaffe *et al.* (2000): conducting a survey on the determinants of citation selection, they discovered that the nature of the technological relationship between two patent could be different. For instance, the citing patent can be seen as an alternative way of doing something that the cited patent did before (i.e. similarity of application); otherwise, it could be that the citing patent does something different than the cited patent, but utilizes a similar method, even though the purpose is different (similarity of technology).

Regarding the impact of the linkages between components, innovation can be achieved through recombining already established elements (Fleming, 2001) or by introducing an established element into a new setting (Hargadon and Sutton, 1997). Fleming (2001) suggested to proxy components with patent technological classes. Multiple classification codes are usually assigned to a patent and they can be used to observe indirectly the process of recombinant search and learning. Such classification codes correspond to well-understood "hardware components" or "building blocks", therefore I derive that their combination can be seen as an architecture.

In summary, backward citations can be considered as a proxy of originality, creativity and novelty of an invention. Moreover, the CPC at level 5 (i.e. the entire code) was used to operationalize the components so that their combination constitutes an architecture. In order to understand whether the combination is "new to the world", I verified if the same combination occurs in the experience period in all the patent applications recorded in PATSTAT database. Only when no identical combination was detected, the focal combination is regarded as new, i.e. the architecture is new.

Therefore, in order to define the type of innovation, the following operationalization was adopted (Figure 11):

- the novelty level of an innovation is analysed through backward citations;
- the impact on the linkages between components is evaluated considering the novelty of the combination of technological fields disclosed in patent documents.

For instance, I define as radical an innovation featured by technological originality and generating a new combination of technology components. Being measured ex-ante, the metric does not take into account the effective market and technological impact of the invention, but only its potential technological innovativeness.

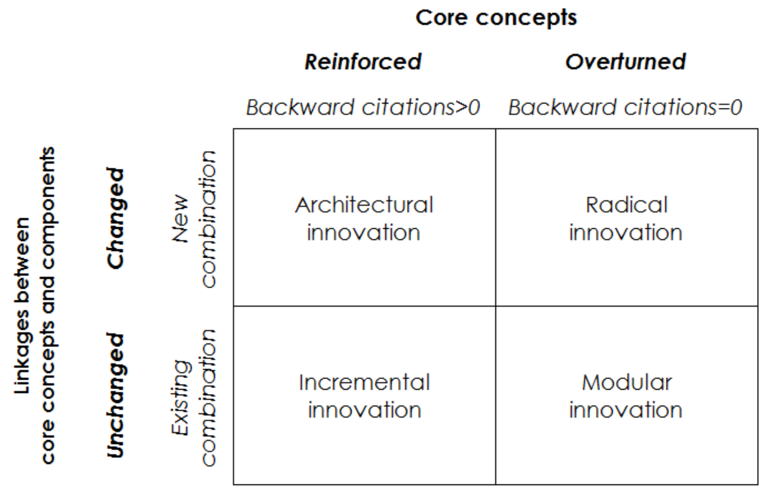

Figure 11. Types of innovation (adapted from Henderson and Clark, 1990).

Therefore, 4 dummy variables can be defined:

- INCR, equal to 1 for incremental innovations;
- MOD, with value 1 if modular innovation has been detected;
- ARCH, assuming 1 if the innovation is architectural;
- RAD, equal to 1 if the innovation is radical.

3.6 Quality of innovation output

The quality of innovation output can be regarded as the impact of the new technology. At a very basic level, a distinction can be made between inventions that impact from a technological perspective (i.e. breakthrough innovation) vs. inventions that are shocks from a user or market perspective (Ahuja and Lampert, 2001).

Regarding breakthrough inventions, they can be identified through patent citation counts. When a patent is cited by other inventors, a higher technical importance can be presumed: patents that are cited are more relevant, innovative and important than those patents that are disregarded (Albert *et al.*, 1991; Rosenkopf and Nerkar, 2001; Dahlin and Behrens, 2005; Alcácer and Gittelman, 2006; Miller *et al.*, 2007; Lahiri, 2010; Mazzucato and Tancioni, 2012). Fleming and Sorenson (2004) argued that the number of forward citations of a patent highly correlates with its technological importance, as measured by expert opinions, social value, and industry awards.

From their survey, Jaffe *et al.* (2000) argued that a patent is cited when its technology is incorporated in other products or new markets are discovered for the same technology. Furthermore, highly cited patents usually lead to higher economic profits than patents that are less frequently cited (Harhoff *et al.*, 1999; Hall *et al.*, 2005). While the number of backward citations is a backward looking measure, which captures the relationship between a patent and the body of knowledge that preceded it, the number of forward citations is a forward-looking measure, which captures the relationship between a patent and subsequent technological developments that build up on it (Mazzucato and Tancioni, 2012).

In this work, I employed information provided by five-year forward citations (Hall *et al.*, 2000; Lahiri, 2010), estimated as the number of patent applications declaring the focal patent in the five years following its publication. Such measure can be regarded as a proxy of the technological value of the invention (Miller *et al.*, 2007).

Yet, not only technological quality has to be considered, since also patents with a low technological impact can have a market impact. Hence, I took into account the patent family size, estimated as the number of all patent applications - priority and non-priority - declaring the focal patent in the five years following its application.

Actually, scholars measure technology marketability by investigating the patent family size. The Organisation for Economic Co-operation and Development (OECD) defines patent families as "the set of patents filed in several countries which are related to each other by one or several priority filings". Given the territorial character of patent protection, when an applicant wants to protect an invention internationally, a patent application has to be filed in each of the countries where protection is sought (either one by one or collectively through supranational filing procedures). As a result, the first patent filing made to protect the invention, the so-called priority filing, which is usually made in the home country of the applicant, is followed by a series of subsequent filings and forms, together with them, a patent family. The number of family patents represents the number of different nations in which a patent is published (Breitzman and Mogee, 2002; Harhoff *et al.*, 2003) and has been considered as indicating the level of R&D or technological activity relevant to international diffusion, thus implying marketability (Geum *et al.*, 2013). Ernst (2003) used the average number of family patents granted by the organization in the technology areas of interest compared to the industry mean for evaluating the firm's products marketability. The family size increases when inventions have been applied for protection in multiple countries (Johnstone *et al.*, 2012), claiming the priority patent. Moreover, follow-on patents, also called continuations, may be further filed by applicants for numerous reasons: to apply for different uses from the same claims, and to file new claims that emerge over time in the R&D process (Gittelman, 2008). Graham and Mowery (2004) reported that about one quarter of patent applications are continuations on prior inventions. They are particularly important in fields where development processes are long, specifically bio-pharmaceuticals. In fields where life cycles are short (e.g., semiconductors) they are more likely to be used for strategic purposes (Graham and Mowery, 2004). Even though many scholars discovered that patents in large families will be more likely to receive higher numbers of future citations (Cockburn and Henderson, 1998; Gittleman and Kogut, 2003; Harhoff *et al.*, 2003), and then technological and market impact are closely related, as mentioned before not all marketable innovations are necessary technological breakthroughs.

Finally, considering that it is expensive for owners to renew patent protection for additional years, I labelled as lapsed a patent whose fees result as unpaid within the eighth year, active otherwise. This can be viewed as a proxy of the value attributed to the invention (Pakes and Simpson, 1989).

Actually, the quality of innovation output and, then, its impact, is defined through ex-post information provided by eight-year renewal fees, five-year forward citations and family size. Specifically, four levels of quality have been introduced (Figure 12):

- low, for all lapsed patents and for active ones when neither forward citations nor patent extension were found;
- technological, when only forward citations were detected;
- market, if only an increase of the family size was discovered;
- high, if both forward citations and patent extensions were uncovered.

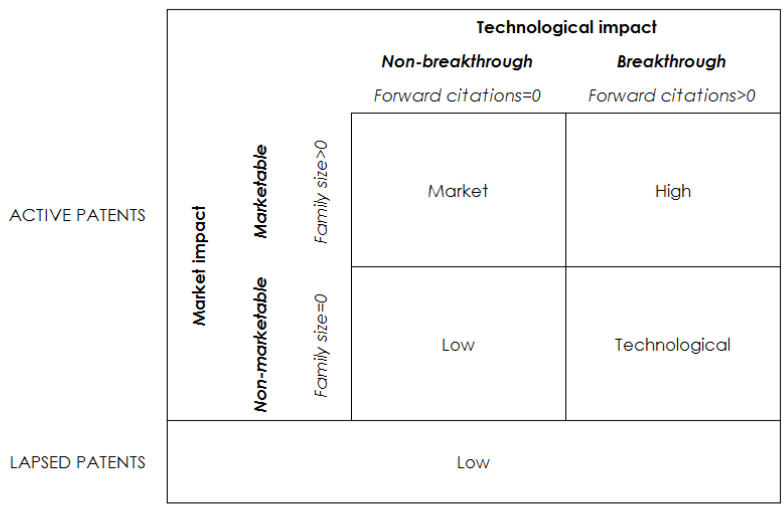

Figure 12. Quality of innovation output.

Therefore, I considered 4 dummy variables (MARK, HIGH, TECH, LOW) capturing the quality of innovation output for each patent application. For instance, when the market quality has been detected, MARK is equal to 1.

3.7 The integrated patent-based framework

The thesis investigates innovation strategies carried out by top R&D spending bio-pharmaceutical and technology hardware & equipment companies ranked by The 2011 EU Industrial R&D Investment Scoreboard. I selected these industries since they are the first for R&D investments and use patents as a means of appropriation of innovation (Pavitt, 1984). I employed information disclosed in patent documents and examined five dimensions of innovation.

The selection of such dimensions has followed an input-process-output-outcome logic. Actually, technological strategy and technological specialization affect the input of the innovative efforts, defining which expertise, capabilities and resources are involved in the R&D process. The "process" block can be summarized by the organizational dimension, which opposes closed, internal development to open, joint development. The type of innovation describes the innovation output, whose quality represents its outcome. Figure 13 displays the innovation dimensions and their variables, also exhibiting the specific patent data employed for the operationalization of each block. The analysis is performed for each patent document; hence, I investigated the five dimensions of innovation within the focal invention.

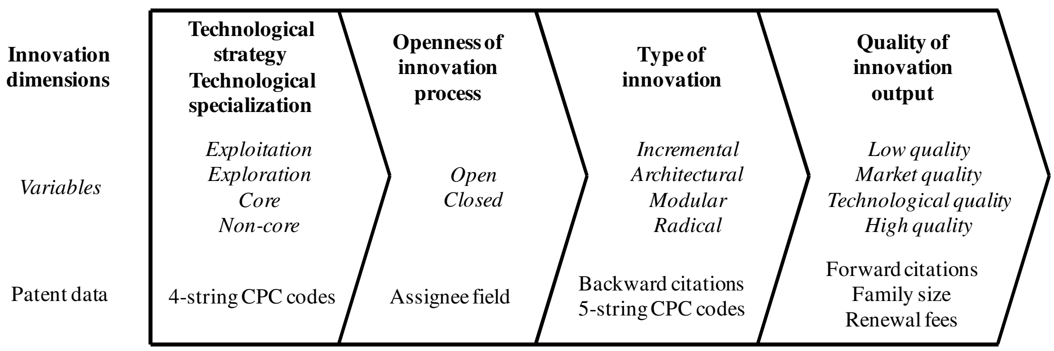

Figure 13. The input-process-output-outcome logic for defining innovation dimensions.

The individual information collected for each patent is used to study the overall behaviour of a firm, cumulating the results obtained from all the documents.

For instance:

- CORE is the share of patenting activities within CPCs labelled as core compared to the total amount of patenting activities in which the firm is involved in t;
- EXPLOR is the share of explorative activities (i.e. activities within CPCs labelled as explorative) compared to the total amount of patent applications in t.

Regarding metrics defined through dummy variables, I divided their occurrence by the total number of patent filed by the focal firm, e.g. RAD is the share of patents labelled as radical on the total number of patent applications owned by the company.

The framework supports in identifying firms' innovation strategies in a specific time interval and provides a useful instrument for benchmarking (i.e. firm-level analysis). Further, by selecting a sample of companies and cumulating the results obtained for each one, the framework also provides information about technological innovation in specific industries, enabling to perform an industry-level analysis.

In Appendix A, I explain the methodology in detail, describing the process of data collection and the programming code developed in order to query the PATSTAT database. I also provide an example of the framework application to patent documents filed by Zeltia, a pharmaceutical firm, from 2003 to 2015. Limitations pertaining to both the use of patent data and the operationalization are outlined in conclusions.

4. FRAMEWORK APPLICATION

The devised framework was applied to a sample of 223 R&D intense companies from BP and THE industries (Appendix B), ranked by their investment in R&D, according to The 2011 EU Industrial R&D Investment Scoreboard (JRC, 2011), in order to perform an industry-level analysis. Firms whose 2011 annual reports were not available on the internet and those for which the list of subsidiaries was not found in such documents were excluded. In detail, the sample consists of 78 BP companies and 145 THE ones. I downloaded from PATSTAT database about 3,000 priority patents filed from 2003 to 2005 by BP companies and more than 20,000 ones applied in the same period from THE ones, performing an industry-level analysis for each sector. Then, I compared the results in order to detect differences in the adoption of innovation strategies between the two industries. Since a study of innovative activities within the experience period is required in order to label knowledge domains and detect technological and specialization strategies, I downloaded about 16,000 patents filed by BP companies before 2003 and about 56,000 applied by THE ones. Each industry-level analysis is performed cumulating the results obtained for each company within the belonging sector.

The data were used from a cross-section perspective, as three years are insufficient for a longitudinal study, especially in an industry like the bio-pharmaceutical where the development time horizon can be longer than ten years. Thus, ANOVA, descriptive statistics and correlation analyses were presented on a set of about 23,000 patent applications and 223 firms.

Since very different business models feature the two industries under investigation, cross-section one-way ANOVA analysis was performed to determine whether the belongingness to the two industries is a discriminating factor for the variables under study (Table 4). All of them, except for the technological quality (TECH) resulted in statistically different mean values between the two industries and, for this reason, both descriptive statistics and correlation analyses were performed and presented separately for the two samples.

Variable		Sum of Squares	df	Mean Square	F	Sig.
	Between	13.196	1	13.196		
CORE	Within	3,334.850	23,331	0.143	92.323	0.000
	Total	3,348.046	23,332			
	Between	13.188	1	13.188		
NONCORE	Within	3,334.883	23,331	0.143	92.266	0.000
	Total	3,348.072	23,332			
	Between	0.920	1	0.920		
EXPLOIT	Within	3,540.668	23,331	0.152	6.064	0.014
	Total	35,41.588	23,332			
	Between	0.916	1	0.916		
EXPLOR	Within	3,540.815	23,331	0.152	6.033	0.014
	Total	3,541.731	23,332			
	Between	1.182	1	1.182		
CLOSED	Within	338.716	23,331	0.015	81.445	0.000
	Total	339.899	23,332			
	Between	0.400	1	0.400		
OPEN_SCI	Within	64.194	23,331	0.003	145.554	0.000
	Total	64.594	23,332			
	Between	0.207	1	0.207		
OPEN_IND	Within	274.888	23,331	0.012	17.536	0.000
	Total	275.095	23,332			
	Between	49.473	1	49.473		
INCR	Within	5,659.991	23,331	0.243	203.930	0.000
	Total	5,709.464	23,332			
	Between	52.919	1	52.919		
MOD	Within	1,369.802	23,331	0.059	901.334	0.000
	Total	1,422.721	23,332			
	Between	17.392	1	17.392		
ARCH	Within	5,748.732	23,331	0.246	70.584	0.000
	Total	5,766.124	23,332			
	Between	15.441	1	15.441		
RAD	Within	1,326.042	23,331	0.057	271.675	0.000
	Total	1,341.483	23,332			
	Between	59.500	1	59.500		
LOW	Within	1,478.112	23,331	0.063	939.165	0.000
	Total	1,537.612	23,332			
	Between	0.00	1	0.000		
TECH	Within	214.982	23,331	0.009	0.022	0.882
	Total	214.982	23,332			
	Between	30.498	1	30.498		
MARK	Within	1,949.889	23,331	0.084	364.920	0.000
	Total	1,980.387	23,332			
HIGH	Between	175.572	1	175.572		
	Within	3,176.022	23,331	0.136	1,289.748	0.000
	Total	3,351.594	23,332			

Table 4. One-Way ANOVA - Discriminating Factor: Industry Belongingness.

4.1 Industry segments

Within each industry, I also defined specific segments, on the basis of the 4-digit ICB codes disclosed in the Scoreboard, in order to detect behaviours depending on the peculiarities of business activities carried out by firms. For the BP industry two segments are defined:

- pharmaceutical companies (PH), relying on a chemical-based synthetic process to develop small-molecule drugs. They generally have greater flexibility to either carry out those functions entirely in-house or license drugs from other entities, including biotech firms, for additional development. Furthermore, PH companies usually generate sales from products they already have on the market. Even though R&D costs are considerable, sales and marketing expenses are particularly relevant for them, since they are necessary to achieve their sales targets and maximizing profitable returns;
- biotech companies (BIO), which use biotechnology to manufacture drugs, involving the manipulation of microorganisms or biological substances to perform a specific process. Focusing primarily on R&D, they discover novel compounds with processes often lengthy, difficult, and costly. Frequently, they operate at a loss for an extended period, and their R&D costs are typically driven by milestone payments related to collaborations with larger, more established biotech or PH companies. If a compound successfully progresses through the final stage of testing, a BIO firm may prepare for commercial launch or can partner up with another company in exchange for a portion of sales.

Regarding the THE industry, three segments can be considered:

- computer hardware & office equipment (CHOE), featured by businesses involved in designing and manufacturing computer hardware and components, such as monitors, data storage, hard drive disks, printers, photocopiers and computer networking infrastructures. The competition among CHOE companies is particularly intense. In the traditional PC market, companies' products have largely become commodified, with constant downward price pressure (and narrowing profit margins) being the result. On the other hand, there are markets for innovative new products, like tablet PCs and ultra-minimal desktops, which are not yet fully commodified. Thus, the race is on to develop products at breakneck speed so they can be first to market. And if a company falters, it instantly becomes a target for larger companies looking to acquire new businesses;

- semiconductors (SE), represented by companies engaged in design and fabrication of semiconductor devices, such as digital and analog integrated circuits. Firms within such segment are technology enablers for the whole electronics value chain. The segment is featured by the need for high degrees of flexibility and innovation in order to constantly adjust to the rapid pace of change in the market. Many products embedding semiconductor devices often have a very short life cycle. At the same time, the rate of constant price-performance improvement in the semiconductor industry is shocking. Consequently, changes in the semiconductor market not only occur extremely rapidly but also anticipate changes in industries evolving at a slower pace;
- telecommunication equipment (TCE), concerning businesses involved in designing and manufacturing hardware used for telecommunications. Actually three main categories of products can be defined: public switching equipments (i.e. analogue and digital switches), transmission equipments (e.g. transmission lines, communications satellites) and private equipments, such as mobile phones, modems and routers. Profitability for individual companies is linked to technical innovation and the ability to secure high-volume contracts from large customers. Small companies can be successful if they make highly specialized products. There are large economies of scale in manufacturing standard products, but many products are specialized and produced in small manufacturing plants.

In what follows, descriptive statistics for the five dimensions under investigation are presented, performing both industry- and segment-level analyses. Finally, for each belonging industry I show the results from the correlation analysis with the aim to give evidence of the relationships among the five dimensions of innovation.

4.2 Technological strategy

Table 5 displays the technological strategies carried out by the sample within the knowledge fields involved in R&D activities. By grouping firms by industry, no significant differences emerge: exploitation strategies are strongly preferred, covering about 70% of innovative activities. Even though companies need to combine both exploitative and explorative activities in order to effectively improve their survival chances and performance, the heritage of routines adopted in the past strongly conditions learning opportunities.

Companies tend to develop new knowledge in domains in which they already possess competencies (Teece, 1986), thus preferring to exploit a technological domain rather than exploring a new one. When they understand the need for a new body of knowledge, they start to explore new knowledge domains, preferring those that are close to ones they currently have at their disposal (Dosi, 1982).

Industry	EXPLOIT	EXPLOR
BP	74.14%	25.86%
THE	72.22%	27.28%

Table 5. Technological strategies by industry.

Only when the analysis is performed at the segment level, specific behaviours emerge (Table 6). Indeed, biotech companies are the most prone to exploration, while among THE firms only semiconductors emerge for their involvement in explorative activities. As a matter of fact, the scientific nature of innovation forces BIO companies to explore new technological fields, in order to discover new compounds.

Segment	EXPLOIT	EXPLOR
BIO	57.58%	42.43%
PH	75.62%	24.38%
CHOE	75.99%	24.01%
SE	67.85%	32.15%
TCE	75.85%	24.15%

Table 6. Technological strategies by industry segment.

Regarding SE firms, exploration is required in order to acquire competencies in new industries in which they aim at entering, such as the automotive one, to provide their skills on chips and integrated circuits in a sector in which electronics is the seed for future innovation.

4.3 Technological specialization

As to technological specialization, BP firms mostly tend to concentrate patenting activity within a familiar and crucial technological field (i.e. core knowledge domain). On the other hand, in the THE industry the capability to recombine and integrate pieces of knowledge belonging to different knowledge domains is primarily critical and leads to a higher breadth of technological fields involved in the development, thus, reducing the average relevance of each CPC (Table 7). Since the production in the THE industry often requires electrical and software engineering competencies and the integration with a variety of components, companies may require knowledge on multiple technologies to work effectively with their suppliers (Brusoni *et al.*, 2001). On the contrary, BP companies are involved in very risky R&D processes that are extremely expensive, take a very long time and have high failure rates (Mazzucato and Tancioni, 2012). Therefore, they are forced to conduct a "guided search", typified by more scale economies in R&D and path-dependency (Gambardella, 1995), and concentrate their activities towards skills that are essential for their survival.

Industry	CORE	NONCORE
BP	32.27%	67.74%
THE	24.98%	75.02%

Table 7. Technological specialization by industry.

At the segment level, the focalization on core knowledge fields is much evident for biotech companies (Table 8).

Segment	CORE	NONCORE
BIO	60.32%	39.70%
PH	29.76%	70.24%
CHOE	20.39%	79.61%
SE	30.78%	69.22%
TCE	19.89%	80.11%

Table 8. Technological specialization by industry segment.

Indeed, in such segment companies have to face many challenges, such as R&D, capital, business development, manufacturing and sales issues, which force them to concentrate their R&D efforts on few knowledge fields. As to the THE industry, SE firms exhibit higher levels of specialization. Such strategy depends on the technological complexity of semiconductors, which forces them to focus mainly on few domains and carry out a specialization strategy. Regarding PH, CHOE and TCE firms, they tend to pursue a diversification strategy, extending their business on a wide range of technological fields. Actually, they mostly develop, manufacture and market many product lines and categories.

4.4 Openness of the innovation process

The closed approach is the strategy more frequently adopted and, in both industries, covers almost totally R&D efforts (Table 9). For instance, in the THE sector companies choose a CI strategy to speed up their R&D processes, being development pace faster and product life cycles shorter. R&D collaborations in such industry cover a small share of innovative activities and are pursued for two main purposes:

- collaborations with other companies, such as SE ones, are necessary since partners manufacture parts, components and products that are incorporated into firms' products and, then, the joint development may improve the overall quality and innovativeness perceived by their customers. Indeed, among THE firms semiconductors display the higher level of R&D collaboration(Table 10);
- R&D partnerships with external firms are stipulated in order to set regulations and industry standards for a particular technology. For instance, it is the case of the computer hardware, where firms define together standards for ports, interfaces, specifications, network architectures, data channels, platform modules.

Regarding companies belonging to the BP industry, they exhibit the higher level of openness, since the integral nature of innovation that features the industry forces them to outsource shares of long, risky and expensive activities and collaborate in order to reduce the overall R&D effort. As a matter of fact, since no single firm possesses all the knowledge, skills and techniques required (Powell *et al.*, 1996), the collaboration results from the need for complementary expertise. The BP industry is featured by the highest rate of joint patenting activities (Kim and Song, 2007) and OI is seen not only as an innovation strategy, but also as the very core business model for many companies, especially for biotech ones. Indeed, most of them do not sell products, but rather enter into collaboration agreements with other BP companies. Such behaviour features specifically small biotech firms collaborating with large pharmaceutical ones.

Industry	CLOSED	OPEN_SCI	OPEN_IND
BP	96,60%	1,40%	2,00%
THE	98,79%	0,13%	1,09%

Table 9. Open innovation adoption by industry

Segment	CLOSED	OPEN_SCI	OPEN_IND
BIO	93.10%	3.23%	3.66%
PH	96.92%	1.23%	1.85%
CHOE	99.42%	0.12%	0.47%
SE	97.82%	0.21%	1.97%
TCE	99.73%	0.01%	0.25%

Table 10. Open innovation adoption by industry segment.

A final remark concerns the adoption of R&D collaboration strategies with scientific partners. In both industries, I uncovered a higher propensity to industrial partnerships, probably because of the features of the investigated sectors. However, BIO companies exhibit the highest level of collaboration with scientific entities. Indeed, they tend to show strong reliance on new discoveries and adaptations made in specialized labs and hence are likely to show high R&D collaboration. Medical innovations have based extensively on activities made possible by interdisciplinary flows of knowledge with life and physical sciences playing key roles. In many cases, BIO companies originate from knowledge spillovers and discoveries within universities and research centres; therefore, they have a natural attitude towards partnerships with scientific entities. In the BP industry, the reason for collaboration with R&D scientific partners include:

- transfer of technology;
- technological/consulting advice;
- technological information absorption;
- access to information on engineers, scientist and trends in R&D;
- research contracts to complement firm R&D;
- research that the firm cannot perform;
- graduate recruitment for supporting R&D activities;
- use of other scientific resources;
- product/process testing;
- quality control improvement.

Regarding THE firms, since they are engaged in standardized product assembly and fast R&D, the large learning process required for partnerships with scientific entities and the fact that the results obtained are, in many cases, not directly exploitable for business applications reduce the propensity to such kind of collaboration.

4.5 Type of innovation

From the combination of the novelty level of the technology and the impact on the linkages between components, I defined for each patent the type of innovation inherent in the application. Table 11 displays the shares of such types of innovation detected within the documents filed by the companies belonging to the sample. In both industries, architectural innovations are most frequent.

Industry	INCR	MOD	ARCH	RAD
BP	30.31%	19.35%	37.28%	13.05%
THE	44.43%	4.75%	45.65%	5.17%

Table 11. Types of innovation by industry.

Since I investigated only priority patents, a new patent family is in many cases originated from the creation of new combinations of components and concepts, without using new disruptive technologies. The second type of innovation that features both industries is the incremental one, more markedly among THE firms. Then, considering that incremental and architectural innovations are based on the reinforcement of core concepts, I deduce that such strategy is the most pursued in these two science-based sectors. As to modular and radical innovation, I discovered the higher propensity to such types of innovation among BP firms.

Given that the nature of products developed by such industries is opposite (i.e. integral in the BP sector and modular in the THE one), I further investigate how firms achieve each type of innovation, in order to validate my operationalization and give an industry-specific interpretation to each one. For the BP industry, I started from the "new drug classification" provided by the Food and Drug Administration (FDA). Actually, FDA labels new drugs along two dimensions at the time of approval: therapeutical potential and chemical composition. On the basis of their therapeutical potential, drugs are classified into two classes: "priority review" drugs, which represent a therapeutical advance over available therapy, and "standard review" drugs, which have therapeutical qualities similar to those of an already marketed drug. Based on their chemical composition, drugs are classified as either new molecular entities (NMEs) or "old" drugs that either are new formulations or have new indications of use. The NMEs are the most technologically advanced products, because they are based on an active ingredient that has never been marketed before. Conceptually speaking, the two dimensions of the FDA classification coincide precisely with the two dimensions in my classification of types of innovation (Figure 14). Specifically, the FDA's therapeutical potential dimension corresponds to the "core concepts" dimension and the chemical composition dimension corresponds to "linkages between components" one (i.e. a new chemical composition can be seen as a new architecture). Therefore, a radical innovation involves products with a substantially new composition and appears to represent an advance over available therapy. A modular innovation provides significantly greater benefits, but the composition is not considerably new. An architectural innovation uses a substantially different composition than existing products but appears to have therapeutical qualities similar to those of an already marketed drug. An incremental innovation consists only on the improvement of a specific drug in terms of limited advancements in safety and efficacy, but BP firms depend on incremental innovations to provide the revenue that will support the development of more risky radical drugs.

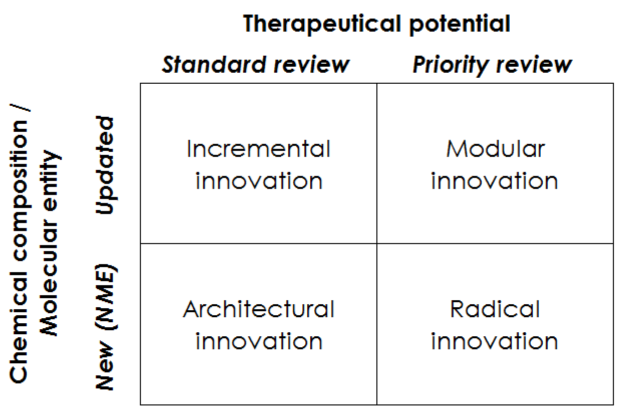

Figure 14. Types of innovation in the BP industry.

As to the THE industry, the labelling is simpler, being products featured by modularity of design. Actually, firms assemble modules, components and parts in order to build the product/hardware architecture. Innovation is then achieved employing two kinds of knowledge, the component knowledge and the architectural one (Figure 15). Therefore:

- incremental innovation refines and extends an established hardware design. Improvement occurs in individual components, but the architecture remain the same;
- modular innovation changes, replaces components without changing the product's architecture. Therefore, the innovation relates to the introduction of new component knowledge in already existing architectures, since the technological novelty is concentrated to these parts;
- architectural innovation is the reconfiguration of an established hardware system to link together existing components in a new way. This does not mean that the components themselves are untouched by architectural innovation, since a change in a component (i.e. size or some other subsidiary parameters) that creates new interactions and new linkages with other components in the established technology may be required. The important point is that the component knowledge, and the associated scientific and engineering knowledge, remain the same;
- radical innovation establishes a new design and, hence, a new component knowledge embodied in hardware components that are linked together in a new architecture. Thus, such type of innovation destroys the usefulness of both previous architectural and component knowledge, which become obsolete.

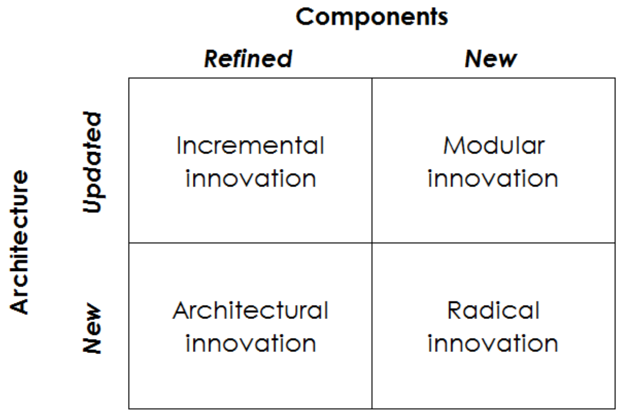

Figure 15. Types of innovation in the THE industry.

A further remark regards radical innovation, since it is the most difficult type of innovation to achieve. Considering the results at the segment level (Table 12), it is possible to appreciate the differences among companies belonging to different segments. In the BP industry, BIO firms emerge for their capability to obtain radical innovations, which is innate in their business. Among THE companies, TCE ones exhibit the higher shares of RAD, thus signalling the vivacity that has always characterized the segment. In general, TCE firms are more prone to architectural innovations than other segments, while CHOE companies are featured by incremental innovation strategies. Such behaviour is still evident in an industry where new dominant designs are difficult to establish and companies are focused on achieving economies of scale in production. Finally, the segment showing the highest share of modular innovations is the pharmaceutical, with firms obtaining an advance over available therapy with already existing compositions.

Segment	INCR	MOD	ARCH	RAD
BIO	30.17%	15.95%	37.07%	16.81%
PH	30.33%	19.65%	37.30%	12.72%
CHOE	51.21%	4.20%	41.32%	3.28%
SE	46.23%	4.27%	45.08%	4.41%
TCE	37.35%	5.80%	49.35%	7.50%

Table 12. Types of innovation by industry segment.

4.6 Quality of innovation output

Given that only the potential quality of a technology is embedded within the concept of type of innovation, the last dimension (i.e. quality of innovation output) investigates the effective contribution of new inventions to companies' business in terms of technological acknowledgement and marketability. Table 13 shows the quality of innovation output by industry: for THE firms I uncovered a larger share of HIGH patents (i.e. patent exhibiting both market and technological quality) compared to BP firms. Specifically, such behaviour can be interpreted with these considerations about the BP industry:

- since a patent is filed before the drug is subject to the evaluation of public health authorities and clinical tests have proved its therapeutical quality, the likelihood of further abandonment is higher. Therefore, in the BP industry the share of lapsed patent affects the quality I detected;
- since MARK is higher in the BP industry, for many patents high quality has been not achieved because of the lack of technological acknowledgement. This is in line with the integral nature of products developed in such sector. Each compound is covered by few patents, each patent protect a "formula" (i.e. a composition), therefore its' very hard to be further cited and technologically acknowledged by third parties. Furthermore, it is simpler for inventions within the THE industry, which are related to application science and engineering, be useful for the technological progress and serve as the basis for subsequent technological development, even in other industries.

Industry	LOW	TECH	MARK	HIGH
BP	20.69%	0.96%	19.10%	59.25%
THE	5.22%	0.93%	8.02%	85.83%

Table 13. Quality of innovation output by industry.

Investigating the quality of innovation output by industry segment (Table 14), I uncovered that CHOE firms exhibit the highest rates of HIGH; therefore, patents in such segment are featured by marketability and technological acknowledgment. Among firms within the BP industry, PH ones show the larger share of low quality applications, while MARK is relevant for BIO firms, thus signalling that in the segment is vital to create a new patent family and aim at the international diffusion.

Segment	LOW	TECH	MARK	HIGH
BIO	15.09%	2.16%	21.98%	60.78%
PH	21.19%	0.85%	18.84%	59.11%
CHOE	1.44%	0.76%	9.34%	88.46%
SE	3.90%	0.29%	9.22%	86.59%
TCE	9.60%	1.93%	5.46%	83.01%

Table 14. Quality of innovation output by industry segment.

4.7 Correlation analyses

The purpose of this section is to show the results of correlation analyses performed for the two industries. The aim is to investigate the relationships among the variables summarising the five dimension of innovation taken into account in my thesis. First, it is necessary to underline that innovation is a complex topic, which is affected by many issues and can be considered as the result of business strategies involving many units, functions and management decisions. Therefore, it is not possible with few variables and five dimensions define the perfect "formula" for successful innovations. The purpose of the section is to delineate which behaviours are related to best outputs and outcomes in each industry under investigation.

Regarding the BP industry, Pearson's correlation coefficients among the variables under investigation in Table 15 indicate that:

- high quality innovations are positively correlated to architectural innovations and negatively to modular ones, therefore investing on new compositions that have therapeutical qualities similar to those of already marketed drugs is the best strategy for successful innovations;
- incremental inventions negatively correlate with low quality outcomes, therefore focusing only on improvements on existing solutions will reduce the likelihood of failure;
- opening up the R&D processes and collaborating with industrial partners lead to architectural innovations, which are acknowledged by the scientific community but not by the market. On the contrary, closed innovation strategies negatively affect the technological quality of an invention;
- among technological strategies, exploitation ones are positively correlated to architectural innovations, therefore firms exploit already owned knowledge in order to develop new compositions;
- companies pursuing a focalization strategy are more prone to open up their R&D processes and collaborate with

industrial partners, providing their specialization and accessing to complementary expertise;

- the focalization is negatively correlated to incremental inventions, therefore firms pursuing a specialization strategy are trying to develop technologies of great potential.

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. CORE	1	-1.000**	.417**	-.417**	-0.148	-0.039	.251*	-.253*	0.183	0.141	-0.051	-0.132	0.18	0.109	-0.023
2. NONCORE		1	-.417**	.417**	0.148	0.039	-.251*	.253*	-0.183	-0.141	0.051	0.132	-0.18	-0.109	0.023
3. EXPLOIT			1	-1.000**	-0.056	0.05	0.028	-0.167	0.058	.228*	-0.098	-0.126	0.164	0.04	0.032
4. EXPLOR				1	0.056	-0.05	-0.028	0.167	-0.058	-.228*	0.098	0.126	-0.164	-0.04	-0.032
5. CLOSED					1	-.712**	-.678**	-0.016	-0.022	-0.134	0.19	0.196	-.455**	-0.046	0.009
6. OPEN_SCI						1	-0.033	0.087	0.12	-0.139	-0.116	-0.159	-0.026	0.153	0.042
7. OPEN_IND							1	-0.068	-0.094	.337**	-0.149	-0.113	.674**	-0.094	-0.057
8. INCR								1	-.524**	-0.196	-.430**	-.255*	-0.116	0.156	0.165
9. MOD									1	-.427**	-0.127	.340**	-0.112	0.041	-.323**
10. ARCH										1	-.265*	-.260*	.361**	-.250*	.322**
11. RAD											1	0.177	-0.09	0.01	-0.147
12. LOW												1	-0.079	-.368**	-.650**
13. TECH													1	-0.134	-0.174
14. MARK														1	-.388**
15. HIGH															1

Note: ** the correlation is significant at 0.01 level, * the correlation is significant at 0.05 level

Table 15. Pearson's correlation coefficients for the BP industry.

As to the THE industry (Table 16), I uncovered:

- a positive relationship between high quality and focalization strategies, thus signalling that firms not dispersing their efforts on a wide range of knowledge fields are likely to achieve successful innovations;
- a positive correlation between high quality and exploitation strategies, therefore avoiding to explore technological domains in which firms lack of prior familiarity will lead to best outcomes;
- that closed innovation strategies correlate with high quality, thus firms gain competitive advantage by internally developing their technologies;
- that radical innovations are positively correlated to focalization strategies, i.e. firms concentrating their efforts develop high-potential technologies which depart significantly from past practices;
- a negative relationship between R&D collaboration with industrial partners and high quality, signalling that firms are unable to employ the results obtained from the joint development for improve their business. Actually, R&D collaborations need time to build up and generate long-term benefits, therefore aren't suitable in industries featured by fast development pace, shorter product life cycles and rapid changes in the market;
- a positive relationship between modular innovations and marketability, thus the overturning of some components within a consolidated architecture lead to new opportunities of development for "old" patent families that are not useful for the scientific community;
- a positive correlation between architectural innovation and technological acknowledgement, signalling that new hardware architectures are further adopted and have an utility on the path of the technological progress, even though the new patent family is no further "enlarged".

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1. CORE	1	-1.000**	.469**	-.469**	0.029	-0.008	-0.028	0.024	-0.121	-0.004	.176*	-.176*	-0.113	-0.134	.248**
2. NONCORE		1	-.469**	.469**	-0.029	0.008	0.028	-0.024	0.121	0.004	-.176*	.176*	0.113	0.134	-.248**
3. EXPLOIT			1	-1.000**	0.162	-0.118	-0.142	0.074	-0.128	-0.024	0.099	-.214**	-0.083	-0.027	.214**
4. EXPLOR				1	-0.162	0.118	0.142	-0.074	0.128	0.024	-0.099	.214**	0.083	0.027	-.214**
5. CLOSED					1	-.304**	-.975**	-0.115	0.102	0.082	-0.101	-.281**	-0.001	-0.049	.251**
6. OPEN_SCI						1	0.086	0.055	-0.058	-0.043	0.078	0.077	-0.011	-0.046	-0.035
7. OPEN_IND							1	0.108	-0.093	-0.076	0.088	.276**	0.003	0.062	-.255**
8. INCR								1	-.299**	-.720**	-.264**	-0.062	-0.148	-0.011	0.105
9. MOD									1	-.376**	0.103	-0.098	-0.046	.281**	-0.049
10. ARCH										1	-0.129	0.156	.195*	-0.143	-0.117
11. RAD											1	-0.096	-0.055	-0.077	0.136
12. LOW												1	0.007	-0.011	-.801**
13. TECH													1	-0.045	-.318**
14. MARK														1	-.487**
15. HIGH															1

Note: ** the correlation is significant at 0.01 level, * the correlation is significant at 0.05 level

Table 16. Pearson's correlation coefficients for the THE industry.

4.8 Discussions

My PhD thesis aims at contributing to the debate on knowledge and innovation management by suggesting a framework for analysing five dimensions of innovation. First, the study of innovation strategies requires some peculiarities:

- the use of standardized data, such as patent documents, allowing to access information about the whole population of innovating firms, since without data regarding all inventions it is unable to absolutely define, for instance, the novelty level of technologies, components and architectures;
- the use of continuously updated data, like patents, since all inventions require time to be acknowledged and adopted by the scientific community and other inventors, therefore *ex-post* information are necessary;
- the assurance that the information has been validated by third parties, like patent examiners, allowing to consider patent data as objective;
- the need for a multidimensional perspective, since R&D development and innovation processes are complex and require a wider overview;
- the necessity to take into account the peculiarities of R&D processes of the industry under investigation, e.g. adapting some operationalisations of variables, and, in addition, the necessity to separate the analysis if more than one industry is studied;

From the analysis of the results, it appears that some are in line with the scientific literature, while others are industry-specific.

By investigating exploitation vs. exploration strategies, I evaluated the share of each strategy on the overall innovation strategy pursued by companies, confirming that exploitation strongly prevails on exploration (Dosi, 1982; Teece, 1986). Even though companies are ambidextrous and the two strategies can coexist inside them (March, 1991; Levinthal and

March, 1993; March, 1996; Tushman and O'Reilly, 1996; Benner and Tushman, 2003; He and Wong, 2004; March, 2006), they cannot within a specific knowledge domain (Burgelman, 2002; Gupta *et al.*, 2006).

Regarding technological specialization, in both industries I found a positive correlation between specialization and exploitation. This confirms that companies at least progressively modify their innovation strategy, avoiding to concentrate a relevant share of R&D activities on previously unexplored domains. In general, industry-specific features, as already discussed in this work, affect the specialization vs. differentiation strategy.

As to OI strategies, the highest levels of openness were detected for BIO firms (about 7%), in line with the real features of the industry segment. In general, the BP sector exhibits the higher share of OI adoption, as already uncovered by scientific literature (Kim and Song, 2007). Regarding the partner typology, no specific behaviours related to partnerships with scientific entities were uncovered, while R&D collaborations with industrial partners have positive effects only in the BP industry.

Regarding the type of innovation, an effort for the industry-specific interpretation of the Henderson and Clark (1990) matrix has proved the application of my operationalization to both sectors. Even though in the analysed industries I uncovered the preponderance of architectural innovations, specific behaviours and correlations emerge. For instance, BP firms are more prone to radical and modular innovations, while THE ones mostly focus on incremental inventions.

As to the quality of innovation output, for THE firms it is simpler to achieve high quality inventions, while in the BP industry many difficulties emerge. Since the economic value of each successful patent in the BP sector is clearly higher, the likelihood of success for a patent is lower.

Finally, correlation analyses demonstrate the effect of industry-specific features on innovation strategies adopted. Only two correlations were confirmed for both industries (i.e. the positive correlation between

CORE and EXPLOIT and the positive relationship between ARCH and TECH). Radical innovations are related to focalization strategies in the THE industry, while no correlation was uncovered for BP firms. Furthermore, high quality innovations are correlated with core, exploitation and closed strategies in the THE sector, and only with architectural innovation in the BP one.

5. CONCLUSIONS

In this final chapter, I present an overview of the theoretical and managerial implications of the work, as well as of the limitations and the future research lines and challenges.

5.1 General contributions

I aim at contributing to the current debate on knowledge and innovation management by providing a patent-based framework, which detects five dimensions of innovation processes and describes how companies organize R&D activities and achieve high quality inventions from a quantitative point of view.

Since innovation processes are featured by extreme complexity, I believe that only multidimensional approaches may better summarize innovative behaviours. The analysis is performed at the knowledge domain level, since the variables under investigation derive from both direct information disclosed in patent documents and in depth studies on technological domains declared in such documents.

I draw on objective data gathered from PATSTAT database and started from variables already acknowledged and operationalized by scholars, improving, adapting and mixing them. As widely described in Appendix A, this work is the result of a detailed study of PATSTAT database, which consists of more than 20 tables and about 500 Gigabytes of data. The methodology and the operationalisation I defined were implemented in a software, which is currently in use at the University of Salerno.

An industry-level analysis on a sample of about 23,000 patent applications and 223 R&D intense companies from BP and THE industries was performed, considering patent applications from 2003 to 2005, validating both the framework applicability and its explicative power and usefulness. Many differences in the adoption of innovation strategies were found, in line with the scientific literature and the characteristics of each industry.

In what follows, the main contributions of the thesis are summarized, distinguishing between implications for theory and practice.

5.2 Theoretical contributions

In this work, I performed a wide literature review aiming at defining which dimensions of innovation were mainly investigated by scholars. Considering the large number of papers and contributions in the field of knowledge and innovation management, I tried to summarise their findings in order to define for each dimension opposite concepts to consider in my framework (e.g. exploitation vs. exploration strategy for the dimension "technological strategy").

Furthermore, from the literature review on the operationalisations mainly acknowledged by scholars I found the relationship between patent data and each dimension of innovation. I critically analysed such metrics and derived new variables from patent data by improving, adapting and mixing those previously adopted in literature.

Hence, the former theoretical contribution concerns the definition of an integrated patent-based framework, which investigates five dimensions of innovation in order to define innovation strategies carried out by firms. Despite innovation strategies are widely studied in literature, most attention has been devoted to only one dimension of R&D processes at time. Since innovation processes are featured by extreme complexity, I suggest a multidimensional approach, which provides a more complete overview of such processes.

Additional theoretical contribution derive from the analysis of specific dimensions. For instance, a contribution pertains to the definition of radical innovation. In my opinion, a distinction between potential radicalness of an invention and real breakthrough is necessary: actually, I derive information about the output of the R&D process from the concept of "type of innovation", with companies aiming, for instance, at achieving a radical innovation. Hence, the metric takes into account *ex-ante* information declared by the company during the application, i.e. defining a potential radicalness. Conversely, the real outcome of the patent can be better measured, according to Arts *et al.* (2013), through *ex-post* indicators, which reflect the impact and value of inventions.

Regarding the types of innovation, my work contributes to the understanding of the difference between radical vs. incremental innovation and exploration vs. exploitation strategy. During my literature review, I uncovered that some scholars use indiscriminately the concept of radical innovation and exploration strategy. Actually, in these works scholars underline that the exploration strategy will lead to radical innovation, while by exploiting the knowledge already owned firms obtain incremental innovation. My work demonstrates that there is no direct relationship. This means that also exploitation strategies may result in radical outputs. For instance, in the BP industry biotech firms employ the few knowledge fields they control in order to obtain potentially novel solutions.

Furthermore, unlike other scholars analyzing exploitation vs. exploration activities, I decided to modify the value of the experience period (i.e. 5 years), since I believe that it is industry-specific. By considering a time span adjustment factor I take into account the different features of the belonging sector of companies and such assumption affects the labelling activity of each technological field owned by firms.

A further theoretical contribution concerns the focus on technological specialization. Actually, I found few contributions regarding specialization vs. differentiation strategies, even though the familiarity of a knowledge field is a key element for R&D processes.

Regarding the openness dimension, I contributed to the wide literature on open innovation by demonstrating that patent data can be useful for detecting R&D collaborations among firms. Therefore, joint patents can be employed for determining the weight and the impact of OI adoption on the overall innovation strategy pursued by firms.

As to the four types of innovation, I believe that the operationalization provided can be considered an interesting contribution, since it recovers the acknowledged matrix provided by Henderson and Clark in 1990 and defines two variables deriving from patent data in order to schematize the novelty level and the linkage between concepts and components. Since such levels are investigated through patent data, the focal patent is really compared to all the other patent inventions in order to define, for instance, whether it is effectively potentially and universally radical. An interesting contribution regards the operationalization of components: I used Cooperative Patent Classification codes in order to define concepts and components and their combination to schematize an architecture.

Finally, regarding the quality of innovation output, unlike many literature contributions, I defined the quality of innovation output as a combination of technological and market impact. Therefore, I believe that the matrix I developed may be the basis for future investigations.

5.3 Practical contributions

I provide a methodology investigating innovation strategies on the basis of the study of patent applications, by using information disclosed in data recorded in PATSTAT database. The advantages in employing patent data in my study are various:

- they are a direct outcome of R&D efforts, and of those inventions which firms expect may have a commercial impact and provide benefits that outweigh costs for obtaining intellectual property protection;
- they contain highly detailed information on content and ownership of patented technology;
- they cover a broad range of technologies.

Given the availability and objectivity of patent documents, studying innovation through the analysis of patent data can help decision-makers to:

- assess the status of firms' innovation strategies;
- monitor the innovation strategy of the company, tracing both its history and evolution within each of the five dimensions. In this way, managers can deal with the challenge of adoption of specific behaviours in their organizations, and make direct improvements, evaluating firm's performance in relation to each one;
- know which innovation-related items to manipulate to improve the effectiveness of innovation strategies;
- compare innovation strategies over time and space;
- benchmark with competitors. Given the availability and objectivity of patent data, the framework can be used as a method of comparability, enabling managers to position their organizations against competitors through a benchmark of innovation strategies.

Moreover, the methodology may be applied at patent-, firm- and industry-level analyses, providing business analysts a practical instrument for detecting the innovation strategies carried out by companies and investigating the impact of specific behaviors on economic and financial performances.

5.4 Limitations

Some limitations can be outlined for the work. Firstly, the use of patenting information as a proxy of inventive activities might underestimate the phenomenon, since not all R&D efforts will result in an application for a patent. Secondly, the research is confined to top R&D spending companies; hence, the findings may not provide a general overview of BP and THE industry as a whole, even though the sample under investigation covers a relevant share of patenting activities within the belonging sector. Thirdly, the use of patent data for investigating the adoption of OI could be questionable, since not all R&D collaborations can be captured by co-patenting activities (Hagedoorn *et al.*, 2003); this may lead to the underestimation of OI activities. Finally, not all technological inventions are patented and patent propensities vary across firms and industries, even though in sectors characterized by intense R&D efforts, like BP and THE ones, patents are used as a means of appropriation of innovation (Pavitt, 1984). This leads to the consideration that my framework may not be useful for analysing innovation in all industries.

Other limitations are related to the operationalization of patent information stored in PATSTAT database. For instance:

- as explained in Appendix A, the patents owned by the investigated firms were detected searching on PATSTAT the names of the parent company and their subsidiaries. Therefore, typing mistakes in person fields within PATSTAT impede the linking between applicants and companies belonging to the sample, thus, some patent applications may have been missed;
- some documents are excluded from the analysis since they did not contain a CPC code;
- the results found in the analysis of technological and specialization strategies are affected by the definition of core and non-core technological fields and, in particular, by the decision of cutting CPCs without considering the sub-group number, in order to avoid excessive detail on the definition of the knowledge domains owned by companies.

This last consideration suggests a deepening of the operationalization of knowledge domain level variables, e.g. building a statistical model in order to define for each industry the proper time spans, rather than identifying them through the analysis of the characteristics of R&D processes.

5.5 Future research directions

In this section, I show some open research lines and challenges for future reference. In fact, the concepts and ideas that were adopted and implemented in this thesis can be used as a basis for a variety of aspects of future work.

Firstly, future research will be addressed to widening the sample of investigation and examining different industries. In addition, it is possible to extend the number of dimensions, e.g. adding other variables from data available or derivable from PATSTAT, such as:

- the diversity of technological fields involved in the invention;
- the typology of the business units which developed the new technology (i.e. parent company, subsidiary, previously acquired or merged unit);
- the technological strategies carried out by partners within the knowledge fields involved in R&D collaborations;
- the scientific complexity of the innovation, which can be measured through non-patent literature cited by the focal patent;
- the impact of external inventors coming from industrial or scientific communities;

In addition, by defining a set of CPCs describing a specific industry, it is possible to go beyond the definition of a sample and analyze the whole industry. This word requires only the development of a new programming code.

Furthermore, the results provided by the framework need to be validated through case studies on companies belonging to the sample and compared or correlated with other databases and sources available, such as annual reports and financial data.

Finally, I plan to apply the content analysis to my framework, in order to replace the CPCs with keywords detected from patents' abstracts. This will improve these dimensions:

- technological strategy and technological specialization, since each keyword may define a knowledge domain;
- type of innovation, since the combination of keywords detected from each abstract may be considered as an architecture.

6. REFERENCES

Abraham, B.P. and Moitra, S.D. (2001), "Innovation assessment through patent analysis", *Technovation*, Vol. 21, No. 4, pp. 245-252.

Acs, Z.J. and Audretsch, D.B. (1989), "Patents as a measure of innovative activity", *Kyklos*, Vol. 42, No. 2, pp. 171-180.

Agryres, N.S. and Silverman, B.S. (2004), "R&D, organization structure and the development of corporate technological knowledge", *Strategic Management Journal*, Vol. 25, No. 89, pp. 929-958.

Ahuja, G. and Katila, R. (2001), "Technological acquisitions and the innovation performance of acquiring firms: a longitudinal study", *Strategic Management Journal*, Vol. 22, No. 3, pp. 197-220.

Ahuja, G. and Lampert, C.M. (2001), "Entrepreneurship in the large corporation: a longitudinal study of how established firms create breakthrough inventions", *Strategic Management Journal*, Vol. 22, No. 6, pp. 521-543.

Al-Ashaab, A., Flores, M., Doultsinou, A. and Magyar, A. (2011), "A balanced scorecard for measuring the impact of industry-university collaboration", *Production Planning & Control*, Vol. 22, No. 5-6, pp. 554-570.

Albert, M.B., Avery, D., Narin, F. and McAllister, P. (1991), "Direct validation of citation counts as indicators of industrially important patents", *Research Policy*, Vol. 20, No. 3, pp. 251-259.

Alcácer, J. and Gittelman, M. (2006), "Patent citations as a measure of knowledge flows: the influence of examiner citations", *Review of Economics and Statistics*, Vol. 88, No. 4, pp. 774-779.

Almeida, P. and Kogut, B. (1999), "Localization of knowledge and the mobility of engineers in regional networks", *Management Science*, Vol. 45, No. 7, pp. 905-917.

Anand, B.N. and Khanna, T. (2000), "Do firms learn to create value? The case of alliances", *Strategic Management Journal*, Vol. 21, No. 3, pp. 295-315.

Archibugi, D. and Pianta, M. (1996), "Measuring technological change through patents and innovation surveys", *Technovation*, Vol. 16, No. 9, pp. 451-468.

Argote, L. (1999), *Organizational learning: creating, retaining and transferring knowledge*, Kluwer Academic Publishers, Boston.

Arora, A. and A. Gambardella (1990), "Complementary and external linkages: the strategies of large firms in biotechnology", *Journal of Industrial Economics*, Vol. 38, No. 4, pp. 361–379.

Arora, A. and Gambardella, A. (1994), "Evaluating technological information and utilizing it", *Journal of Economic Behavior and Organization*, Vol. 24, No. 1, pp. 91-114.

Arts, S. (2012), "Path dependency and recombinant exploration: how established firms can outperform in the creation of technological breakthroughs", Working paper.

Arts, S. and Veugelers, R. (2012), "The technological origins and novelty of breakthrough inventions", Working paper.

Arts, S., Appio, F.P. and Van Looy, B. (2013), "Inventions shaping technological trajectories: do existing patent indicators provide a comprehensive picture?", *Scientometrics*, Vol. 97, No. 2, pp. 397-419.

Baptista, R. and Mendonça, J. (2010), "Proximity to knowledge sources and the location of knowledge-based start-ups", *The Annals of Regional Science*, Vol. 45, No. 1, pp. 5-29.

Basalla, G. (1988), *The evolution of technology*, Cambridge: Cambridge University Press.

Baum, J.A.C. and C. Oliver (1991), "Institutional linkages and organizational mortality", *Administrative Science Quarterly*, Vol. 36, No. 2, pp. 187-218.

Bayona Saez, C., Garcia Marco, T. and Huerta Arribas, E. (2002), "Collaboration in R&D with universities and research centres: an empirical study of Spanish firms", *R&D Management*, Vol. 32, No. 4, pp. 321-341.

Belderbos, R., Carree, M. and Lokshin, B. (2004), "Cooperative R&D and firm performance", *Research Policy*, Vol. 33, No. 10, pp. 1477-1492.

Belderbos, R., Faems, D., Leten, B. and Van Looy, B. (2010), "Technological activities and their impact on the financial performance of the firm: exploitation and exploration within and between firms", *Journal of Product Innovation Management*, Vol. 27, No. 6, pp. 869-882.

Benfratello, L. and Sembenelli, A. (2002), "Research joint ventures and firm level performance", *Research Policy*, Vol. 31, No. 4, pp. 493-507.

Benner, M.J. and Tushman, M. (2002), "Process management and technological innovation: a longitudinal study of the photography and paint industries", *Administrative Science Quarterly*, Vol. 47, No. 4, pp. 676-706.

Benner, M.J. and Tushman, M. (2003), "Exploitation, exploration, and process management: the productivity dilemma revisited", *Academy of Management Review*, Vol. 28, No. 2, pp. 238-256.

Breitzman, A.F. and Mogee, M.E. (2002), "The many applications of patent analysis", *Journal of Information Science*, Vol. 28, No. 3, pp. 187-205.

Brouthers, K.D., Brouthers, L.E. and Wilkinson, T.J. (1995), "Strategic alliances: choose your partners", *Long Range Planning*, Vol. 28, No. 3, pp. 18-25.

Brusoni, S., Prencipe, A. and Pavitt, K. (2001), "Knowledge, specialization, organizational coupling, and the boundaries of the firm: why do firms know more than they make?", *Administrative Science Quarterly*, Vol. 46, No. 4, pp. 597-621.

Burgelman, R.A. (2002), “Strategy as vector and the inertia of coevolutionary lock-in”, *Administrative Science Quarterly*, Vol. 47, No. 2, pp. 325-357.

Chakrabarti, A.K. (1991), “Competition in high technology: analysis of patents of US, Japan, UK, West Germany, and Canada”, *IEEE Transactions of Engineering Management*, Vol. 38, No.1, pp. 78-84.

Chang, S.J. (1996), “An evolutionary perspective on diversification and corporate restructuring: entry, exit, and economic performance during 1981-89”, *Strategic Management Journal*, Vol. 17, No. 8, pp. 587-611.

Chesbrough, H. (2003), *Open Innovation: the new imperative for creating and profiting from technology*, Boston: Harvard Business School Press.

Chesbrough, H. (2006), *Open business models: how to thrive in the new innovation landscape*, Boston: Harvard Business School Press.

Christensen, C.M. and Overdof, M. (2000), “Meeting the challenge of disruptive change”, *Harward Business Review*, Vol. 78, No. 2, pp. 66–76.

Cockburn, I.M. and Henderson, R.M. (1998), “Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery”, *The Journal of Industrial Economics*, Vol. 46, No. 2, pp. 157-182.

Cohen, W.M. and Levinthal, D.A. (1990), “Absorptive capacity: a new perspective on learning and innovation”, *Administrative Science Quarterly*, Vol. 35, No. 1, pp. 128-152.

Cyert, R.M. and Goodman, P.S. (1997), “Creating effective university-industry alliances: an organizational perspective”, *Organizational Dynamics*, Vol. 5, No. 4, pp. 45-57.

Dahlin, K.B. and Behrens, D.M. (2005), "When is an invention really radical? Defining and measuring technological radicalness", *Research Policy*, Vol. 34, No. 5, pp. 717-737.

Das, T.K. and Teng, B.S. (2000), "A resource-based theory of strategic alliances", *Journal of Management*, Vol. 26, No. 1, pp. 31-61.

David, P.A. and Steinmueller, W.E. (1995), "Economics of compatibility standards and competition in telecommunication networks", *Information Economics and Policy*, Vol. 6, No. 3–4, pp. 217-242.

Dittrich, K. and Duysters, G. (2007), "Networking as a means to strategy change: the case of open innovation in mobile telephony", *Journal of Product Innovation Management*, Vol. 24, No. 6, pp. 510-521.

Dosi, G. (1982), "Technological paradigms and technological trajectories", *Research Policy*, Vol. 11, No. 3, pp. 147-162.

Duysters, G. and de Man, A.P. (2003), "Transitory alliances: an instrument for surviving turbulent industries?", *R&D Management*, Vol. 33, No. 1, pp. 49-57.

Egyedi, T.M. (1999), "Examining the relevance of paradigms to base OSI standardisation", *Computer Standards & Interfaces*, Vol. 20, No. 4, pp. 355-374.

Ernst, H. (2003), "Patent information for strategic technology management", *World Patent Information*, Vol. 5, No. 3, pp. 233-242.

Fleming, L. (1998), *Explaining the source and tempo of invention: recombinant learning and exhaustion in technological evolution*, Stanford: Stanford University Press.

Fleming, L. (2001), "Recombinant uncertainty in technological search", *Management Science*, Vol. 47, No. 1, pp. 117-132.

Fleming, L. and Sorenson, O. (2004), "Science as a map in technological search", *Strategic Management Journal*, Vol. 25, No. 89, pp. 909-928.

Gambardella, A. (1995), *Science and innovation in the US Pharmaceutical Industry*, Cambridge University Press, Cambridge.

Geum, Y., Lee, S., Yoon, B. and Park, Y. (2013), "Identifying and evaluating strategic partners for collaborative R&D: index-based approach using patents and publications", *Technovation*, Vol. 3, No. 6-7, pp. 211-224.

Gittelman, M. (2008), "A note on the value of patents as indicators of innovation: implications for management research", *Academy of Management Perspectives*, Vol. 22, No. 3, pp. 21-27.

Gittelman, M. and Kogut, B. (2003), "Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns", *Management Science*, Vol. 49, No. 4, pp. 366-382.

Graff, G.D. (2003), "Observing technological trajectories in patent data: empirical methods to study the emergence and growth of new technologies", *American Journal of Agricultural Economics*, Vol. 85, No. 5, pp. 1266-1274.

Graham, S. and Mowery, D. (2004), "Submarines in software? Continuations in U.S. software patenting in the 1980s and 1990s", *Economics of Innovation and New Technology*, Vol. 13, No. 5, pp. 443-456.

Granovetter, M.S. (1973), "The strength of weak ties", *American Journal of Sociology*, Vol. 78, No. 6, pp. 1360-1380.

Griliches, Z. (1990), "Patent statistics as economic indicators: a survey", *Journal of Economic Literature*, Vol. 28, No. 4, pp. 1661-1707.

Grupp, H. (1992), *Dynamics of science-based innovation*, Springer, Berlin.

Gupta, A.K., Smith, K.G. and Shalley, C.E. (2006), "The interplay between exploration and exploitation", *Academy of Management Journal*, Vol. 49, No. 4, pp. 693-706.

Hagedoorn, J., van Kranenburg, H. and Osborn, R. N. (2003), "Joint patenting amongst companies - exploring the effects of inter-firm R&D partnering and experience", *Managerial and Decision Economics*, Vol. 24, No. 2-2, pp. 71-84.

Hall, B., Jaffe, A. and Trajtenberg, M. (2000b), "Market value and patent citations: a first look", NBER working paper.

Hall, B., Jaffe, A. and Trajtenberg, M. (2001), "The NBER patent citations data file: lessons, insights and methodological tools", NBER working paper.

Hall, B., Jaffe, A. and Trajtenberg, M. (2005), "Market value and patent citations", *RAND Journal of Economics*, Vol. 36, No. 1, pp. 16-38.

Hall, B., Link, A. and Scott, J. (2000a), "Universities as research partners", working paper, National Bureau of Economic Research.

Hargadon, A. (2003), *How breakthroughs happen: the surprising truth about how companies innovate*, Boston: Harvard Business School Press.

Hargadon, A. and Sutton, R.I. (1997), "Technology brokering and innovation in a product development firm", *Administrative Science Quarterly*, Vol. 42, No. 4, pp. 716-749.

Harhoff, D., Scherer, F.M. and Vopel, K. (2003), "Citations, family size, opposition and the value of patent rights", *Research Policy*, Vol. 32, No. 8, pp. 1343-1363.

Hawkins, R. (1999), "The rise of consortia in the information and communication technology industries: emerging implications for policy", *Telecommunications Policy*, Vol. 23, No. 2, pp. 159-173.

He, Z.L. and Wong, P.K. (2004), "Exploration vs. exploitation: an empirical test of the ambidexterity hypothesis", *Organization Science*, Vol. 15, No. 4, pp. 481-494.

Henderson, R.M. and Clark, K.B. (1990), "Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms", *Administrative Science Quarterly*, Vol. 35, No. 1, pp. 9-30.

Hung, S. W. and Tang, R. H. (2008), "Factors affecting the choice of technology acquisition mode: an empirical analysis of the electronic firms of Japan, Korea and Taiwan", *Technovation*, Vol. 28, No. 9, pp. 551-563.

Jaffe, A., Trajtenberg, M. and Fogarty, M.S. (2000), "The meaning of patent citations: report on the NBER/case-western reserve survey of patentees", NBER Working Paper.

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993), "Geographic localization of knowledge spillovers as evidenced by patent citations", *Quarterly Journal of Economics*, Vol. 108, No. 3, pp. 577-598.

Johnstone, N., Haščič, I., Poirier, J., Hemar, M. and Michel, C. (2012), "Environmental policy stringency and technological innovation: evidence from survey data and patent counts", *Applied Economics*, Vol. 44, No. 17, pp. 2157-2170.

JRC (2011), "The 2011 EU Industrial R&D Investment Scoreboard", Joint Research Centre, Directorate General Research and Innovation, European Commission.

Kaplan, S. and Vakili, K. (2012), "Identifying breakthroughs: cognitive vs. economic", Working paper, DRUID 2012, Denmark.

Kim, C. and Song, J. (2007), "Creating new technology through alliances: an empirical investigation of joint patents", *Technovation*, Vol. 27, No. 8, pp. 461-470.

Kogut, B. (1988), "Joint ventures: theoretical and empirical perspectives", *Strategic Management Journal*, Vol. 9, No. 4, pp. 319-332.

Koza, M.P. and Lewin, A.Y. (1998), "The co-evolution of strategic alliances", *Organization Science*, Vol. 9, No. 3, pp. 255-264.

Krackhardt, D. (1992), "The strength of strong ties: the importance of philos in organizations", In N. Nohria & R. Eccles (eds.), *Networks and*

Organizations: Structure, Form and Action: 216-239, Harvard Business School Press, Boston.

Lahiri, N. (2010), "Geographic distribution of R&D activity: how does it affect innovation quality?", *Academy of Management Journal*, Vol. 53, No. 5, pp. 1194-1209.

Leonard-Barton, D. (1992), "Core capabilities and core rigidities: a paradox in managing new product development", *Strategic Management Journal*, Vol. 13, No. S1, pp. 111-125.

Leten, B., Belderbos, R. and Van Looy, B. (2007), "Technological diversification, coherence, and performance of firms", *Journal of Product Innovation Management*, Vol. 24, No. 6, pp. 567-579.

Levinthal, D.A. and March, J.G. (1993), "The myopia of learning", *Strategic Management Journal*, Vol. 14, No. S2, pp. 95-112.

Liebeskind, J. P., A. L. Oliver, L. Zucker and M. Brewer (1996), "Social networks, learning and flexibility: Sourcing scientific knowledge in new biotechnology firms", *Organization Science*, Vol. 7, No. 4, pp. 428-443.

Lo Storto, C. (2006), "A method based on patent analysis for the investigation of technological innovation strategies: The European medical prostheses industry", *Technovation*, Vol. 26, No. 8, pp. 932-942.

Ma, Z. and Lee, Y. (2008), Patent application and technological collaboration in inventive activities: 1980-2005, *Technovation*, Vol. 28, No. 6, pp. 379-390.

March, J. (1991), Exploration and exploitation in organizational learning, *Organization Science*, Vol. 2, No. 1, pp. 71-87.

March, J. (1996), "Continuity and change in theories of organizational action", *Organization Science*, Vol. 41, No. 2, pp. 278-287.

March, J. (2006), "Rationality, foolishness, and adaptive intelligence", *Strategic Management Journal*, Vol. 27, No. 2, pp. 201-214.

Mazzucato, M. and Tancioni, M. (2012), "R&D, patents and stock return volatility", *Journal of Evolutionary Economics*, Vol. 22, No. 4, pp. 811-832.

Michelino, F., Lamberti, E., Cammarano, A. and Caputo, M. (2015a), "Open models for innovation: an accounting-based perspective", *International Journal of Technology Management*, Vol. 68, No. 1/2, pp. 99-121.

Michelino, F., Lamberti, E., Cammarano, A. and Caputo, M. (2015b), "Measuring open innovation in the bio-pharmaceutical industry", *Creativity and Innovation Management*, Vol. 24, No. 1, pp. 4-28.

Miller, D. J. (2004), "Firms' technological resources and the performance effects of diversification: a longitudinal study", *Strategic Management Journal*, Vol. 25, No. 11, pp. 1097-1119.

Miller, D.J., Fern, M.J. and Cardinal, L.B. (2007), "The use of knowledge for technological innovation within diversified firms", *Academy of Management Journal*, Vol. 50, No. 2, pp. 308-326.

Miner, A.S., Bassoff, P. and Moorman, C. (2001), "Organizational improvisation and learning: a field study", *Administrative Science Quarterly*, Vol. 46, No. 2, pp. 304-337.

Nakamura, M. (2003), "Research alliances and collaborations: introduction to the special issue", *Managerial and Decision Economics*, Vol. 24 No. 2/3, pp. 47-49.

Nakamura, M., Shaver, J.M. and Yeung, B. (1996), "An empirical investigation of joint venture dynamics: evidence from U.S.-Japan joint ventures", *International Journal of Industrial Organization*, Vol. 14, No. 4, pp. 521-541.

Nayyar, P.R. and Kazanjian, R.K. (1993), "Organizing to attain potential benefits from information asymmetries and economies of scope in related diversified firms", *Academy of Management Review*, Vol. 18, No. 4, pp. 735-759.

Nelson, R. and Winter, S.J. (1982), *An evolutionary theory of economic change*, The Belknap Press of Harvard University Press, Cambridge.

Nerkar, A. and Roberts, P.W. (2004), "Technological and product-market experience and the success of new product introductions in the pharmaceutical industry", *Strategic Management Journal*, Vol. 25, No. 89, pp. 779-799.

Nielsen, B, (2003), "An empirical investigation of the drivers of international strategic alliance formation", *European Management Journal*, Vol. 21, No. 3, pp. 301-322.

O'Reilly, C.A. and Tushman, M.L. (2004), "The ambidextrous organization", *Harvard Business Review*, Vol. 82, No. 1, pp. 74-82.

Okubo, Y. and Sjöberg, C. (2000), "The changing pattern of industrial scientific research collaboration in Sweden", *Research Policy*, Vol. 9, No. 1, pp. 81-98.

Pakes, A. and Simpson, M. (1989), "Patent renewal data, Brookings Papers of Economic Activity", *Microeconomics*, Vol. 5, pp. 331-410.

Patzelt, H. and Shepherd, D.A. (2008), "The decision to persist with underperforming alliances: the role of trust and control", *Journal of Management Studies*, Vol. 45 No. 7, pp. 1217-43.

Pavitt, K. (1984), "Sectorial patterns of technical change: towards a taxonomy and a theory", *Research Policy*, Vol. 13, No. 6, pp. 343-373.

Pavitt, K., Robson, M. and Townsend, J. (1989), "Technological accumulation, diversification and organization in UK companies, 1945-1983", *Management Science*, Vol. 35, No. 1, pp. 81-99.

Perez L., Whitelock, J. and Florin, J. (2013), "Learning about customers", *European Journal of Marketing*, Vol. 47 No. 3/4, pp. 431-462.

Pisano, G.P. (1990), "The R&D boundaries of the firm: an empirical analysis", *Administrative Science Quarterly*, Vol. 35, No. 1, pp. 153-176.

Powell, W.W., Koput, K.W. and Smith-Doerr, L. (1996), "Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology", *Administrative Science Quarterly*, Vol. 41, No. 1, pp. 116-145.

Prahalad, C. and Ramaswamy, V. (2004), *The future of competition: co-creating unique value with customers*, Harvard Business School Press, Boston, MA.

Rausser, G. (1999), "Private/public research: knowledge assets and future scenarios", *American Journal of Agricultural Economics*, Vol. 81, No. 5, pp. 1011-1027.

Rosenkopf, L. and Nerkar, A. (2001), "Beyond local search: boundary-spanning, exploration, and impact in the optical disk industry", *Strategic Management Journal*, Vol. 22, No. 4, pp. 287-306.

Sakakibara, M. (1997), "Evaluating government-sponsored R&D consortia in Japan: who benefits and how?", *Research Policy*, Vol. 26, No. 4/5, pp. 447-473.

Sakata, J., Suzuki, K. and Hosoya, J. (2009), " The analysis of research and development efficiency in Japanese companies in the field of fuel cells using patent data", *R&D Management*, Vol. 39, No. 3, pp. 291-304.

Schmidt, S.K. and Werle, R. (1998), *Coordinating Technology: Studies in the international standardization of telecommunications*, Oxford University Press, New York.

Schroll, A. and Mild, A. (2011), "Open innovation modes and the role of internal R&D: an empirical study on open innovation adoption in Europe", *European Journal of Innovation Management*, Vol. 14, No. 4, pp. 475-495.

Schulz, M. (2001), "The uncertain relevance of newness: organizational learning and knowledge flows", *Academy of Management Journal*, Vol. 44, No. 4, pp. 661-681.

Shan, W., G. Walter and B. Kogut (1994), “Interfirm cooperation and startup innovation in the biotechnology industry”, *Strategic Management Journal*, Vol. 15, No. 5, pp. 387-394.

Shane, S. (2001), “Technological opportunities and new firm creation”, *Management Science*, Vol. 47, No. 2, pp. 205-220.

Simon, H.A. (1991), “Bounded rationality and organizational learning”, *Organization Science*, Vol. 2, No. 1, pp. 125-134.

Stinchcombe, A. L. (1965), “Social structure and organizations”, in J. G. March (ed.), *Handbook of Organizations*, Rand McNally, Chicago, IL, pp. 142-193.

Stuart, T. E. (1998), “Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry”, working paper, Graduate School of Business, University of Chicago.

Suzuki, J. and Kodama, F. (2004), “Technology diversity of persistent innovators in Japan: two case studies of large Japanese firms”, *Research Policy*, Vol. 33, No. 3, pp. 531-549.

Teece, D. J. (1992), “Competition, cooperation, and innovation: organizational arrangements for regimes of rapid technological progress”, *Journal of Economic Behavior and Organization*, Vol. 18, No. 1, pp. 1-25.

Trajtenberg, M. (1990a), *Economic analysis of product innovation: the case of CT scanners*, Harvard University Press, Cambridge.

Trajtenberg, M. (1990b), “A penny for your quotes: patent citations and the value of information”, *RAND Journal of Economics*, Vol. 21, No. 1, pp. 325-342.

Trajtenberg, M., Henderson, R. and Jaffe, A. (1992), “Ivory tower versus corporate lab: an empirical study of basic research and appropriability”, NBER working paper.

Trajtenberg, M., Jaffe, A. and Henderson, R. (1997), "University versus corporate patents: a window on the basicness of invention", *Economics of Innovation and New Technology*, Vol. 5, No. 1, pp. 19-50.

Tsang, E.W.K. (1997), "Choice of international technology transfer mode: a resource-based view", *Management International Review*, Vol. 37, No. 2, pp. 151-168.

Tushman, M.L. and Anderson, P. (1986), "Technological discontinuities and organizational environments", *Administrative Science Quarterly*, Vol. 31, No. 3, pp. 439-465.

Tushman, M.L. and O'Reilly, C.A. (1996), "Ambidextrous organizations: managing evolutionary and revolutionary change", *California Management Review*, Vol. 38, No. 4, pp. 8-30.

Tyler, B.B. and Steensma, H.K. (1995), "Evaluating technological collaborative opportunities: a cognitive modelling perspective", *Strategic Management Journal*, Vol. 16, No. S1, pp. 43-70.

Venaik, S., Midgley, D. E. and Devinney, T. M. (2005), "Dual paths to performance: the impact of global pressures on MCN subsidiary conduct and performance", *Journal of International Business Studies*, Vol. 36, No. 6, pp. 655-675.

Wu, W.Y., Shih, H.A. and Chan, H.C. (2009), "The analytic network process for partner selection criteria in strategic alliances", *Expert Systems with Applications*, Vol. 36, No. 3, pp. 4646-4653.

APPENDIX A. PATSTAT database, data collection and example of framework application

The purpose of this section is to illustrate how information from PATSTAT database has been employed in order to collect data and implement the patent-based framework.

PATSTAT contains bibliographical and legal status patent data relating to more than 90 million patent documents from over 100 countries, collected by the European Patent Office (EPO). This is extracted from the EPO's databases and is provided as raw data for statistical tools. Raw data contain 220 Gigabytes of information, which reach about 500 Gigabytes considering the addition of new tables and variables I made in order to adequate such data to my research purposes and the space dedicated to MySQL database, on which I imported raw data and implemented my tool.

The interface used for the tool is web-based, built on a LAMP platform (Linux, Apache, MySQL and PHP) and the open source software XAMPP for Windows. The programming code was developed using HTML, CSS, JAVA and primarily PHP scripts and a considerable commitment was required to write about 50,000 lines of code used to manage the database and collect data.

In what follows I show an example of data collection and management considering the information extracted for Zeltia, a Spanish pharmaceutical company that operates through a variety of subsidiaries within the bio-pharmaceutical industry. Until 2007, Zeltia's research activities in the pharmaceutical area had not resulted in a marketed product.

The first step of the work is the download of the list of subsidiaries from firm's 2011 annual report. These are the units disclosed in such document:

- Pharma Mar, S.A.U.;
- Genómica, S.A.U.;
- Zelnova, S.A.;
- Protección de Maderas, S.A.U.;

- Xylazel, S.A.;
- Pharma Mar USA;
- Pharma Mar AG (Switzerland);
- Pharma Mar SARL (France);
- Pharma Mar GmbH (Germany);
- Pharma Mar Ltd (UK);
- Copyr, S.p.A. (Italy);
- Promaxsa Protección de Maderas, S.L.U.;
- Sylentis, S.A..

For each subsidiary, I searched more information, in order to verify the occurrence of units merged or acquired by other companies. Actually, only patent applications filed after the merge/acquisition were considered in the analysis, therefore I recorded the eventual year of the merge/acquisition event. In the case of Zeltia, all the units are direct subsidiaries established by the company group itself.

For each subsidiary a research on the table *tls206_person* in PATSTAT was performed, aiming at finding the records containing the name of each unit in the field *person_name* and resulting in the following list of distinct names:

- GENOMICA S.A.U;
- GENOMICA S.A.U.;
- GENOMICA S.A.V;
- NOSCIRA S.A.;
- NOSCIRA, S. A.;
- NOSCIRA, S.A.;
- PHARMA MAR;
- PHARMA MAR S. A.;
- PHARMA MAR S. A. SOC UNIPERSONAL;
- PHARMA MAR S. A. U.;
- PHARMA MAR S.A;
- PHARMA MAR S.A SOCIEDAD UNIPERSONAL;
- Pharma Mar S.A.;
- PHARMA MAR S.A. SOCIEDAD UNIPERSONAL;
- Pharma Mar S.A., Sociedad Unipersonal;
- PHARMA MAR S.A., SOCIEDAD UNIPERSONAL.;
- PHARMA MAR S.A.SOCIEDAD UNIPERSONAL;
- Pharma Mar S.A.U.;
- PHARMA MAR S.A.U., COLMENAR VIEJO;
- PHARMA MAR SA;

- PHARMA MAR SA SOC UNIPERSONAL;
- PHARMA MAR SAU;
- PHARMA MAR, S;
- PHARMA MAR, S. A. U.;
- PHARMA MAR, S.A;
- PHARMA MAR, S.A.;
- Pharma Mar, S.A. Sociedad Unipersonal;
- Pharma Mar, S.A., a Madrid, Spain corporation;
- PHARMA MAR, S.A., SOCIEDAD UNIPERSONAL;
- PHARMA MAR, S.A..U.;
- Pharma Mar, S.A.U.;
- PHARMA MAR,S.A.;
- PharmaMar;
- PharmaMar, s.a.;
- SYLENTIS S. A.;
- Sylentis S.A.;
- Sylentis S.A.U..

The example provides a clear understanding about the lack of unique correspondence between the name of a subsidiary and the records in the table *tls206_person*. Actually, many records were found for each subsidiary, and occasionally the names are affected by typing errors, since EPO recorded information in PATSTAT database with automatic procedures, which detect text through optical character recognition (OCR) software. Nevertheless, often PATSTAT relates such *person_name* field to a standardized name reported in table *tls208_doc_std_nms*, therefore each standardized name is linked to a wide number of records in the table *tls206_person*, which can be rapidly assigned to the firm. Actually, the process starts with the search within the standardized names and ends with a further research among the residual records, which are not yet linked to the company.

At the end of the process, all the "person names" related to Zeltia were recorded and assigned with the firm. Thus, a first list of patent applications can be defined by extracting data from the table *tls201_appln*. The list consists of 360 patent applications from 2003 to 2005, but only 8 documents are related to priority patent applications filed in EPO, United States Patent Office (USPTO) or WIPO (World Intellectual Property Organization). The other documents are:

- subsequent filings aiming at extending the protection in other countries;
- application continuations to pursue additional claims to an invention already filed in a parent application;
- priority filings not recorded in EPO, USPTO or WIPO, excluded since information recorded in other offices lack of some data, such as citations;
- applications without CPCs recorded in the table *tls224_appln_cpc*, which joins the table *tls201_appln* – containing patent applications – with the technological classes affected by the invention;
- subsequent filings necessary to complete the process of approval and provide further information to examiners.

The 8 priority patents can be detected by linking the information from table *tls201_appln* with the priority applications disclosed in table *tls219_inpadoc_fam*, which stores the list of all patent families and their parent application (Table 17).

Application ID	Family ID	Application number	Filing date	#CPCs
209115	143104	EP20030779140	20/10/2003	4
192977	143104	EP20040714847	26/02/2004	4
16138696	311134	EP20040720081	12/03/2004	1
16161426	742363	EP20040768394	09/09/2004	4
16173267	744877	EP20040798717	15/11/2004	4
16205256	334715	EP20050075779	04/04/2005	2
16206407	341717	EP20050077333	12/10/2005	8
16270510	1850300	EP20050803151	26/10/2005	5

Table 17. List of Zeltia's patent applications under investigation.

This is a sample of the SQL query used for extracting the previous table:

```
SELECT distinct tls201_appln.appln_id, tls219_inpadoc_fam.inpadoc_family_id,
tls201_appln.appln_id .appln_nr, tls201_appln. appln_filing_date, appln_ncpc
FROM tls201_appln, tls207_pers_appln, tls206_person, tls219_inpadoc_fam,
tls204_appln_prior
WHERE tls204_appln_prior.prior_appln_id=tls201_appln.appln_id and
tls206_person.person_id=tls207_pers_appln.person_id and
tls201_appln.appln_id=tls207_pers_appln.appln_id and
tls206_person.companyID=$IDcompany and
tls219_inpadoc_fam.appln_id=tls201_appln.appln_id and
tls207_pers_appln.applt_seq_nr>0 and tls207_pers_appln.applt_fake=0 and
appln_ncpc>0 and yearacquisition_person<= applyear and (appln_auth='WO' or
appln_auth='EP' or appln_auth='US') and applyear between 2003 and 2005
ORDER BY tls201_appln. applyear, tls201_appln.appln_id

(the PHP variable $ IDcompany includes an identification number assigned to Zeltia
and stored in the table tls206_person in order to create a relationships between the
table and a list of companies under investigation)
```

Therefore, taking into account only priority documents that meet the previous features I reach the final list of 8 inventions developed by Zeltia from 2003 to 2005.

For each document the following information are collected:

- number of backward citations, by querying the table *tls212_citation*, with the focal patent recorded as the citing document. From the count are excluded non-patent literature citations;
- number of 5-year forward citations, as the number of patent applications citing the focal patent in the five years following its publication, by investigating the table *tls212_citation*;
- number of patent assignees, as the number of distinct applicants recorded in the patent application. Various procedures are required in order to avoid corrupted counts. First, if the same company has been recorded more than one time within the field (e.g. the patent reports both the parent company and a subsidiary), the applicants have been considered as an unique entity. Second, in many cases in the applicant field were wrongly disclosed inventor names. Therefore, I developed an algorithm for removing

"fake applicants" from the list, which verify if the item has been simultaneously recorded in both applicant and inventor lists and remove the inventors from the assignee field;

- family size, as the number of patents declaring the focal filing as claimed priority in the five years following its application.

In addition to these variables, other data are extracted from PATSTAT, whose operationalization is much complex. The first variable is "new combo", which assumes value 1 if the combination of CPCs declared in the patent application can be considered as new (i.e. not disclosed in previous documents), otherwise 0. Each CPCs (at level 5) is collected from the table *tls224_appln_cpc*. In order to limit the years of investigation and avoid the comparison with obsolete technologies, I verified if the same combination occurs in the experience period in all patent applications recorded in PATSTAT database. For instance, considering that Zeltia belongs to the bio-pharmaceutical industry, I took into account all filings from t-1 to t-7, where t is the year of application of the focal patent. In order to limit the computational impact of such an extended analysis, if the number of CPCs recorded in the focal application is less than or equal to 2 the variable "new combo" assumes automatically value 0, supposing that a document disclosing at least 2 classification codes may not cover an invention based on a novel architecture.

A further variable, which requires a specific operationalization, is the expiration date of the patent. Actually, in my framework I verify if the fees result paid within the eighth year from the first application. In order to detect the expiration date I analyse the table *tls221_inpadoc_prs* containing legal status information (i.e. *ex-post* events that impact on the patent). More specifically, the expiration date is the minimum value deriving from three searches I made within such table. The first queries the field *l513ep*, which directly contains the expiration date, but it is not used by all the patent offices. The second verifies the content recorded in the notes to an event. If specific keywords, such as "LAPSED" or

"EXPIRED" are reported within the field, I extract the event date and assume that it matches with the expiration one. This is a sample of the SQL query used for the purpose:

```
SELECT min(year(`l525ep`)),min(year(`prs_gazette_date`))
FROM tls221_inpadoc_prs
WHERE `appln_id`=$patentID and (`l510ep` LIKE '%LAPSED IN%' or `l510ep`  LIKE '%LAPSE
DUE TO%' or `l510ep`  LIKE '%LAPSE BECAUSE%' or `l510ep`  LIKE '%LAPSE AS%' or `l510ep`
LIKE '%LAPSE/EXPIRED%' or `l510ep`  LIKE '%LAPSED THROUGH%' or `l510ep`  LIKE '%HAS
LAPSED%' or `l510ep`  LIKE '%APPLICATION LAPSED%' or `l510ep`  LIKE '%APPLICATIONS
LAPSED%' or `l510ep`  LIKE '%EXPIRED ON%' or `l510ep` LIKE '%EXPIRED IN%' or `l510ep`
LIKE '%WITHDRAW%')

($patentID is the variable containing an identifier of the focal patent)
```

The third value is derived by employing the field *l520ep* within the legal status in which PATSTAT counts the years from the patent grant. If such value is equal to 20, the date of the event is collected. Finally, by comparing these three dates, I consider the minimum value as the reference date for the expiration.

The software I developed automatically launches such operationalisations. Therefore, the user has only to link the records from the table "person" with the names of the subsidiaries owned by the company under investigation and the following processes were automatically executed by the application. A second activity that users have to perform regards the definition of the partner typology. Indeed, a list of partners disclosed in patent applications under investigation is proposed to the user, who has to define if the partner is an industrial or a scientific entity. For the patents filed by Zeltia, two documents disclose a joint development with third parties. More specifically, the partners are:

- DANA-FARBER CANCER INSTITUTE, INC., a scientific organization which conducts community-based programs in cancer prevention, detection, and control;
- ORTHO BIOTECH PRODUCTS L.P., a biotech company (i.e. industrial entity) acquired by J&J in 2008.

Considering that the number of investigated partners may be significant, I also developed the following function, which suggests the partner typology, but leaves the decision to the user.

```
function scientific_person($partner) {
$partner=strtoupper($partner);
if (
strpos($partner,'COLLEGE') !== false || strpos($partner,'UNIVERSITY') !== false ||
strpos($partner,'UNIVERISTY') !== false || strpos($partner,'UNIVERSTY') !== false ||
strpos($partner,'MEDICAL CENTER') !== false || strpos($partner,'UNIVERSITAET') !== false
|| strpos($partner,'UNIVERSITAT') !== false || strpos($partner,'CLINIC') !== false ||
strpos($partner,'INSTITUTE') !== false || strpos($partner,'INSTITUT') !== false ||
strpos($partner,'INSTYTUT') !== false || strpos($partner,'HOSPITAL') !== false ||
strpos($partner,'FOUNDATION') !== false || strpos($partner,'EDUCATION') !== false ||
(strpos($partner,'RESEARCH') !== false && strpos($partner,' INC') === false) ||
(strpos($partner,'RESEARCH') !== false && strpos($partner,' LTD') === false) ||
strpos($partner,'COUNCIL') !== false ||
(strpos($partner,'NATIONAL') !== false && strpos($partner,'INTERNATIONAL') === false) ||
(strpos($partner,'INST') !== false && strpos($partner,'SCIENCE') !== false) ||
(strpos($partner,'INST') !== false && strpos($partner,'TECH') !== false) ||
(strpos($partner,'INST') !== false && strpos($partner,'RES') !== false) ||
(strpos($partner,'INST') !== false && strpos($partner,'NAT') !== false) ||
strpos($partner,'INST OF') !== false || strpos($partner,'DEPARTMENT') !== false ||
strpos($partner,'MINISTER FOR') !== false || strpos($partner,'MINISTER OF') !== false ||
strpos($partner,'MINISTRY OF') !== false || strpos($partner,'REGENTS') !== false ||
strpos($partner,'AGENCY') !== false || strpos($partner,'GOVERNMENT') !== false ||
strpos($partner,'GOVT') !== false || strpos($partner,'INTERUNIVERSITAIR') !== false ||
strpos($partner,'INTERUNIVERSITY') !== false || strpos($partner,'RECHERCHE') !== false ||
strpos($partner,'POLYTECHNIC') !== false || strpos($partner,'CENTRUM') !== false ||
strpos($partner,'CENTER FOR THE') !== false || strpos($partner,'UNIVERSITEIT') !== false ||
strpos($partner,'CONSORZIO PER LA RICERCA') !== false ||  strpos($partner,'ISTITUTO DI
RICERCA') !== false || strpos($partner,'UNIVERSIT') !== false || strpos($partner,'ECOLE
SUPERIEURE') !== false || strpos($partner,'CONSIGLIO NAZIONALE') !== false ||
strpos($partner,'COMMISSARIAT') !== false || strpos($partner,'POLITECNICO DI') !== false
|| strpos($partner,'POLITECN') !== false || strpos($partner,'UNIVERSITE') !== false ||
strpos($partner,'UNIVERSIDAD') !== false || strpos($partner,'ASSOCIATION') !== false ||
strpos($partner,'CONSEJO SUPERIOR') !== false || strpos($partner,'ASSOCIAZIONE') !==
false || // strpos($partner,'UNITED STATES') !== false || da solo UNITED STATES non basta
strpos($partner,'JUNAJTED STEHJTS') !== false || strpos($partner,'UNIV ') !== false ||
strpos($partner,' UNIV') !== false || (strpos($partner,'LABORATORY') !== false &&
strpos($partner,' INC') === false) || strpos($partner,'DEPT ') !== false ||
strpos($partner,'JUNIVERSITI') !== false || strpos($partner,'UNIVERSTIY') !== false ||
strpos($partner,'ZENTRUM') !== false  || strpos($partner,'AUTHORITY') !== false  ||
strpos($partner,'THE STATE OF') !== false  || strpos($partner,'SCHOOL') !== false  ||
strpos($partner,'SENTRUM') !== false
)
 { $scientifi=1;}
else { $ scientifi =0;}
return $ scientifi;
}
```

(If at least one of these conditions is verified -i.e. the name contains at least one of these words – the partner is labelled as scientific, otherwise as industrial)

After having completed this labelling, the user can run the script, which will display the results of the analysis for all the dimensions under investigation. Table 18 shows the values of CORE, NONCORE, EXPLOIT and EXPLOR within the 8 patents filed by Zeltia. Each variable exhibits the share of CPCs, which are labelled as core, non-core, exploitative or explorative. For these operationalisations, I used CPCs until the main group, i.e. codes at level 4.

Application number	CORE	NONCORE	EXPLOIT	EXPLOR
EP20030779140	33.00%	67.00%	33.00%	67.00%
EP20040714847	67.00%	33.00%	33.00%	67.00%
EP20040720081	100.00%	0.00%	100.00%	0.00%
EP20040768394	67.00%	33.00%	67.00%	33.00%
EP20040798717	100.00%	0.00%	0.00%	100.00%
EP20050075779	0.00%	100.00%	0.00%	100.00%
EP20050077333	33.00%	67.00%	33.00%	67.00%
EP20050803151	50.00%	50.00%	50.00%	50.00%
Share	**56.25%**	**43.75%**	**39.50%**	**60.50%**

Table 18. Focalization and technological strategies from Zeltia's patents

For the focalization dimension, it is necessary to investigate the share of patent applications filed from t-1 to t-7 (i.e. the experience period) which contain each technological domain.

For instance, for the application EP20030779140, with filing year 2003, the following CPCs were extracted:

- "A61K 38", recorded in 17% of the patent applications between 1996 and 2002, and then, considered as core, since it exceeds the 10% threshold, as defined in section 3.3;
- "C07K 7", reported in 6% of filings from 1996 to 2002, labelled as non-core technological domain;
- "C07K 14", not disclosed in any patent application during the experience period, therefore considered as non-core.

Since only one-third of the domains was labelled as core, CORE is 33%, while NON-CORE is 67%. This means that the invention involves 2 technological fields, which are not relevant for Zeltia, and only 1 on which the firm has focalized many R&D efforts in the experience period. The average focalization (i.e. CORE) detected from Zeltia's patents is measured as the mean value from the 8 investigated patents and is equal to 56.25%. This means that Zeltia is strongly specialized in more than 50% of technological domains recorded in its applications.

Regarding the definition of technological strategies carried out by the firm, a CPC is labelled as exploitative if at least one patent containing such CPC has been recorded in the previous years. As argued in section 3.2, for the bio-pharmaceutical industry the period ranges from t-7 to t-4. Therefore, for the patent EP20030779140 filed in 2003 by Zeltia, the analysis is performed by considering patent applications from 1996 to 1999. Among the three CPCs found in this patent, only "A61K 38" was labelled as exploitative, since 1 patent was filed from 1996 to 1999, while "C07K 7" and "C07K 14" are considered as explorative fields. Therefore, the invention has been developed by exploiting 1 technological field and exploring 2 new knowledge domains, thus EXPLOIT is 33% and EXPLOR 67%. Overall, from 2003 to 2005 Zeltia carries out prevalently an exploration strategy, since it covers the 60.5% of CPCs recorded in the 8 patent documents.

Regarding the dimension related to the "openness of the innovation process", I employ the number of applicants collected for each patent. If the document records only Zeltia in the assignee field, the variable CLOSED is set to 1, otherwise 0. When partners are found within the field, I extract the partner typology: OPEN_SCI is equal to 1 if the partner is a scientific entity, otherwise OPEN_IND is set to 1. Table 19 displays the results from Zeltia's patent applications. The closed innovation strategy covers 75% of the total R&D efforts, while the two shares of open innovation strategies are both equal to 12.5%.

Application number	CLOSED	OPEN_SCI	OPEN_IND
EP20030779140	1	0	0
EP20040714847	1	0	0
EP20040720081	0	1	0
EP20040768394	1	0	0
EP20040798717	1	0	0
EP20050075779	1	0	0
EP20050077333	1	0	0
EP20050803151	0	0	1
Share	**75.00%**	**12.50%**	**12.50%**

Table 19. Openness of the innovation process from Zeltia's patents

As to the type of innovation dimension, I employ information provided by backward citations and "new combo" variables. Actually, following the operationalization described in section 3.5, I detect:

- an incremental innovation, where the patent discloses backward citations and the combination of CPCs was already used by prior art:
- a modular innovation, if no citations to previous contributions were found and the variable "new combo" is equal to 0;
- an architectural innovation, when the technology cites other patents and the combination is new;
- a radical innovation, if no citations to prior art were uncovered and no patents using the same combination were found.

Table 20 summarises the results obtained for Zeltia. Modular and radical innovations are more frequent, and each one covers 37.50% of technologies developed by the firm from 2003 to 2005.

Application number	Backward citations	New combo	INCR	MOD	ARCH	RAD
EP20030779140	0	1	0	0	0	1
EP20040714847	0	0	0	1	0	0
EP20040720081	0	0	0	1	0	0
EP20040768394	0	0	0	1	0	0
EP20040798717	0	1	0	0	0	1
EP20050075779	1	0	1	0	0	0
EP20050077333	3	1	0	0	1	0
EP20050803151	0	1	0	0	0	1
Share			**12.50%**	**37.50%**	**12.50%**	**37.50%**

Table 20. Type of innovation from Zeltia's patents

Regarding the last dimension under investigation (i.e. quality of innovation output), three variables are used in order to define the impact of the technology: expiration date, forward citations and family size. Actually, if the patent is lapsed within the following 8 years from the application, the quality is automatically LOW. It is the case of 2 applications filed by Zeltia (EP20050075779 and EP20050077333). For the other documents, it is necessary check market and technological impact, following the operationalization described in section 3.6. Actually, the patent is labelled as:

- LOW, when neither forward citations nor patent extension were found;

- TECH, if only forward citations were recorded;
- MARK, when only an increase of the family size was uncovered;
- HIGH, if both forward citations and patent extensions were detected.

Table 21 displays the result obtained for the 8 patent filed by Zeltia. The inventions are mainly featured by market quality, except for the two lapsed applications.

Application number	Expiration Date	Forward Citations	Family size	LOW	TECH	MARK	HIGH
EP20030779140	2099	0	25	0	0	1	0
EP20040714847	2099	0	12	0	0	1	0
EP20040720081	2099	0	7	0	0	1	0
EP20040768394	2099	0	5	0	0	1	0
EP20040798717	2099	0	1	0	0	1	0
EP20050075779	2010	1	10	1	0	0	0
EP20050077333	2007	0	23	1	0	0	0
EP20050803151	2099	0	8	0	0	1	0
Share				**25.00%**	**0.00%**	**75.00%**	**0.00%**

Table 21. Quality of innovation output from Zeltia's patents

Therefore, the information collected for Zeltia can be used to perform a firm-level analysis, also to benchmark with competitors. Furthermore, by cumulating all the patent applications filed by a sample of firms it is possible to carry out an industry-level analysis. Indeed, in chapter 4 I present descriptive statistics regarding the whole sample under investigation, in which I have considered 23,000 patent application collected from PATSTAT. In addition, in section 4.7 two correlation analyses were performed in order to define the relationships between the strategies pursued by firms in each dimension.

APPENDIX B. List of companies

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
Abbott	PH	5%	95%	77%	23%	100%	0%	0%	36%	12%	48%	3%	7%	1%	15%	77%
Active	BIO	0%	100%	0%	100%	100%	0%	0%	0%	50%	0%	50%	0%	0%	0%	100%
Adtran	TCE	17%	84%	0%	100%	100%	0%	0%	50%	0%	50%	0%	0%	0%	0%	100%
Advanced Digital Broadcast	SE	0%	100%	0%	100%	100%	0%	0%	33%	0%	67%	0%	0%	0%	67%	33%
Advanced Micro Devices	SE	42%	58%	88%	12%	96%	0%	4%	50%	7%	38%	5%	1%	1%	8%	90%
Advanced Semiconductor Eng.	SE	63%	37%	0%	100%	100%	0%	0%	0%	11%	67%	22%	0%	0%	11%	89%
Advantest	SE	65%	35%	72%	28%	100%	0%	0%	55%	2%	42%	2%	4%	0%	6%	91%
Affymetrix	BIO	48%	52%	73%	27%	100%	0%	0%	14%	43%	14%	29%	0%	0%	14%	86%
Alexion	BIO	59%	41%	0%	100%	100%	0%	0%	25%	25%	38%	13%	0%	0%	38%	63%
Allergan	PH	45%	55%	81%	19%	99%	1%	0%	36%	22%	28%	14%	1%	0%	13%	86%
Almirall	PH	24%	76%	82%	18%	100%	0%	0%	64%	36%	0%	0%	36%	0%	45%	18%
Altera	SE	53%	47%	54%	46%	100%	0%	0%	68%	0%	31%	1%	1%	0%	3%	96%
Amgen	BIO	44%	57%	75%	25%	92%	0%	8%	27%	11%	54%	8%	0%	0%	14%	86%
Amkor Technology	SE	100%	0%	100%	0%	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%
Analog Devices	SE	1%	99%	44%	56%	100%	0%	0%	50%	1%	49%	0%	0%	0%	4%	96%
Apple	CHOE	43%	57%	62%	38%	100%	0%	0%	43%	4%	50%	3%	0%	0%	2%	98%
Applied Materials	SE	52%	48%	86%	14%	100%	0%	0%	31%	12%	43%	14%	5%	0%	5%	90%
ARM	SE	45%	55%	26%	74%	92%	8%	0%	60%	6%	34%	0%	2%	2%	8%	88%
Arris	TCE	17%	83%	79%	21%	100%	0%	0%	26%	12%	47%	15%	2%	1%	4%	93%
ASM International	SE	56%	44%	71%	29%	96%	2%	2%	38%	6%	49%	8%	0%	0%	8%	92%
ASML Holding	SE	77%	23%	82%	18%	98%	0%	2%	57%	5%	34%	4%	1%	0%	11%	88%

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
Asustek Computer	CHOE	50%	50%	0%	100%	50%	9%	42%	50%	0%	0%	50%	0%	0%	0%	100%
Atmel	SE	32%	68%	43%	57%	100%	0%	0%	61%	2%	37%	0%	1%	0%	10%	89%
ATMI	SE	26%	74%	54%	46%	100%	0%	0%	41%	5%	53%	0%	2%	0%	2%	97%
Austriamicrosystems	SE	0%	100%	0%	100%	100%	0%	0%	60%	0%	40%	0%	80%	0%	0%	20%
Avago Technologies	SE	21%	79%	21%	79%	100%	0%	0%	50%	0%	50%	0%	2%	0%	8%	91%
Avaya	TCE	35%	65%	66%	34%	99%	0%	1%	31%	6%	57%	7%	0%	0%	6%	94%
Axis	CHOE	0%	100%	0%	100%	100%	0%	0%	0%	0%	100%	0%	100%	0%	0%	0%
Bavarian	BIO	72%	29%	18%	82%	100%	0%	0%	25%	0%	13%	63%	38%	0%	50%	13%
Biotest	PH	67%	33%	0%	100%	100%	0%	0%	33%	33%	33%	0%	0%	0%	67%	33%
Boehringer	PH	35%	65%	79%	21%	99%	1%	0%	24%	43%	23%	10%	20%	0%	45%	34%
BristolMyers	PH	18%	82%	79%	21%	97%	2%	1%	51%	10%	31%	8%	10%	3%	19%	68%
Broadcom	SE	8%	92%	82%	18%	100%	0%	0%	34%	2%	60%	3%	0%	0%	5%	94%
Brocade Communications S.	TCE	0%	100%	0%	100%	50%	0%	50%	50%	0%	50%	0%	0%	0%	0%	100%
Brother Industries	CHOE	22%	78%	22%	78%	67%	0%	33%	89%	0%	11%	0%	22%	11%	22%	44%
BTG	BIO	14%	86%	7%	93%	100%	0%	0%	13%	13%	13%	63%	50%	0%	13%	38%
Bull	CHOE	67%	33%	100%	0%	100%	0%	0%	100%	0%	0%	0%	0%	0%	67%	33%
Cabot Microelectronics	SE	75%	25%	73%	27%	100%	0%	0%	55%	16%	24%	5%	0%	0%	7%	93%
Cadila Healthcare	PH	16%	84%	0%	100%	100%	0%	0%	0%	33%	0%	67%	100%	0%	0%	0%
Calix	TCE	0%	100%	0%	100%	100%	0%	0%	50%	0%	50%	0%	0%	0%	0%	100%
Canon	CHOE	56%	44%	66%	34%	81%	6%	14%	61%	3%	33%	3%	3%	0%	6%	92%
Celgene	BIO	64%	36%	55%	45%	100%	0%	0%	37%	3%	50%	10%	7%	3%	13%	77%
Chiesi	PH	63%	38%	60%	40%	100%	0%	0%	25%	50%	8%	17%	33%	0%	42%	25%
CHR	PH	4%	96%	2%	98%	100%	0%	0%	50%	0%	25%	25%	25%	0%	25%	50%
Ciena	TCE	43%	57%	65%	35%	100%	0%	0%	64%	0%	36%	0%	0%	0%	0%	100%

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
Cisco Systems	TCE	46%	54%	56%	44%	100%	0%	0%	36%	2%	58%	4%	1%	0%	3%	96%
Comverse Technology	TCE	50%	50%	40%	60%	100%	0%	0%	20%	40%	40%	0%	0%	0%	20%	80%
Corning	TCE	32%	68%	75%	25%	100%	0%	0%	39%	0%	61%	0%	1%	1%	4%	94%
Cray	CHOE	31%	69%	11%	89%	100%	0%	0%	53%	0%	47%	0%	0%	0%	6%	94%
Cree	SE	63%	37%	76%	24%	100%	0%	0%	43%	3%	53%	1%	0%	1%	1%	98%
CSL	BIO	0%	100%	0%	100%	0%	100%	0%	100%	0%	0%	0%	0%	0%	0%	100%
CSR UK	TCE	0%	100%	0%	100%	100%	0%	0%	50%	0%	50%	0%	0%	0%	25%	75%
Cypress Semiconductor	SE	9%	91%	16%	84%	100%	0%	0%	62%	0%	38%	0%	0%	3%	11%	86%
Dell	CHOE	15%	85%	8%	92%	100%	0%	0%	52%	15%	27%	6%	0%	0%	4%	96%
Delta Electronics	CHOE	67%	33%	65%	35%	100%	0%	0%	67%	11%	22%	0%	6%	0%	6%	89%
Dendreon	BIO	100%	0%	0%	100%	50%	0%	50%	50%	50%	0%	0%	0%	0%	50%	50%
Dialog Semiconductor	SE	5%	95%	18%	82%	100%	0%	0%	75%	3%	21%	0%	12%	0%	38%	49%
Diamyd	PH	0%	100%	0%	100%	100%	0%	0%	0%	33%	0%	67%	0%	0%	33%	67%
Diebold	CHOE	50%	50%	0%	100%	100%	0%	0%	40%	0%	60%	0%	0%	0%	0%	100%
Dong-A Pharmaceutical	PH	0%	100%	0%	100%	100%	0%	0%	0%	0%	100%	0%	0%	0%	0%	100%
Dynavax Technologies	BIO	100%	0%	33%	67%	100%	0%	0%	0%	33%	67%	0%	100%	0%	0%	0%
Egis	PH	36%	64%	0%	100%	100%	0%	0%	0%	100%	0%	0%	100%	0%	0%	0%
Eisai	PH	69%	31%	62%	39%	100%	0%	0%	13%	75%	6%	6%	6%	0%	44%	50%
Elan	PH	66%	34%	75%	25%	54%	31%	15%	23%	8%	38%	31%	8%	8%	8%	77%
Electronics for imaging	CHOE	50%	50%	57%	43%	100%	0%	0%	57%	4%	35%	4%	0%	0%	4%	96%
EliLilly	PH	19%	81%	70%	30%	95%	0%	5%	12%	37%	12%	38%	49%	2%	35%	14%
ELMOS Semiconductor	SE	43%	57%	20%	80%	67%	0%	33%	33%	0%	67%	0%	67%	0%	33%	0%
EMC	CHOE	67%	33%	83%	17%	100%	0%	0%	37%	10%	49%	4%	1%	3%	1%	95%
Emulex	CHOE	48%	52%	48%	52%	100%	0%	0%	74%	0%	26%	0%	0%	0%	0%	100%

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
Entegris	SE	34%	66%	36%	64%	97%	0%	3%	45%	0%	55%	0%	0%	0%	5%	95%
Ericsson	TCE	11%	89%	96%	4%	99%	0%	1%	40%	4%	48%	8%	14%	14%	7%	65%
Evotec	PH	23%	77%	28%	72%	60%	0%	40%	50%	0%	50%	0%	20%	0%	20%	60%
Exelixis	PH	8%	92%	8%	92%	100%	0%	0%	33%	0%	50%	17%	50%	0%	17%	33%
Extreme Networks	TCE	36%	64%	7%	93%	100%	0%	0%	42%	0%	58%	0%	0%	0%	0%	100%
F5 Networks	TCE	11%	89%	0%	100%	100%	0%	0%	67%	0%	33%	0%	0%	0%	0%	100%
Fairchild Semiconductor	SE	56%	44%	62%	38%	100%	0%	0%	43%	0%	57%	0%	0%	0%	4%	96%
FEI	SE	63%	38%	48%	52%	95%	0%	5%	14%	0%	81%	5%	5%	0%	0%	95%
Finisar	TCE	61%	39%	66%	34%	100%	0%	0%	53%	5%	41%	1%	0%	0%	6%	94%
Forest	PH	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%
GedeonRichter	PH	29%	71%	29%	71%	100%	0%	0%	0%	88%	0%	13%	75%	0%	25%	0%
Gilead	BIO	58%	42%	53%	47%	86%	9%	5%	10%	24%	41%	24%	7%	0%	52%	41%
Glenmark Pharmaceuticals	PH	0%	100%	0%	100%	100%	0%	0%	0%	100%	0%	0%	100%	0%	0%	0%
GN Store Nord	TCE	28%	72%	50%	50%	100%	0%	0%	33%	0%	50%	17%	0%	0%	17%	83%
Hanmi Pharm	PH	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%
Harmonic	TCE	0%	100%	0%	100%	100%	0%	0%	33%	33%	33%	0%	0%	0%	0%	100%
Harris	TCE	5%	95%	48%	52%	100%	0%	0%	35%	7%	54%	4%	0%	0%	8%	92%
Hewlett-Packard	CHOE	3%	97%	84%	16%	100%	0%	0%	55%	2%	43%	1%	1%	1%	11%	87%
Hikma	PH	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	0%	0%	100%	0%
Hisamitsu Pharmaceutical	PH	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%
HTC	TCE	34%	66%	9%	91%	100%	0%	0%	21%	7%	64%	7%	57%	14%	14%	14%
Huawei Technologies	TCE	0%	100%	0%	100%	100%	0%	0%	0%	0%	100%	0%	50%	0%	0%	50%
Hynix Semiconductor	SE	25%	75%	0%	100%	0%	0%	100%	65%	0%	35%	0%	75%	0%	15%	10%
Infineon Technologies	SE	27%	73%	82%	18%	90%	0%	10%	51%	3%	43%	2%	6%	1%	16%	77%

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
Integrated Device Technology	SE	17%	83%	7%	93%	100%	0%	0%	38%	28%	24%	10%	0%	0%	10%	90%
Intel	SE	0%	100%	85%	15%	100%	0%	0%	45%	3%	49%	4%	0%	0%	5%	95%
Interdigital	SE	41%	59%	70%	30%	100%	0%	0%	24%	9%	53%	14%	0%	0%	6%	93%
Intermec	CHOE	40%	60%	10%	90%	100%	0%	0%	60%	10%	20%	10%	0%	0%	0%	100%
International Rectifier	SE	42%	58%	52%	48%	100%	0%	0%	45%	1%	52%	2%	1%	3%	8%	88%
Intersil	SE	34%	66%	24%	76%	100%	0%	0%	59%	7%	34%	0%	0%	0%	3%	97%
Ipsen	PH	0%	100%	0%	100%	100%	0%	0%	0%	100%	0%	0%	100%	0%	0%	0%
Isis	BIO	83%	17%	93%	7%	96%	4%	0%	37%	7%	48%	7%	4%	0%	19%	78%
J&J	PH	36%	64%	87%	13%	99%	0%	1%	30%	7%	56%	7%	7%	1%	9%	83%
JDS Uniphase	TCE	19%	81%	40%	60%	100%	0%	0%	34%	2%	63%	2%	0%	0%	10%	90%
Juniper Networks	TCE	63%	37%	63%	37%	100%	0%	0%	54%	0%	46%	0%	0%	0%	0%	100%
Kla-Tencor	SE	55%	45%	74%	26%	100%	0%	0%	45%	1%	54%	0%	0%	1%	6%	93%
Krka	PH	49%	51%	0%	100%	100%	0%	0%	48%	28%	24%	0%	38%	0%	21%	41%
Kulicke & Soffa	SE	65%	35%	63%	37%	100%	0%	0%	8%	8%	83%	0%	0%	0%	8%	92%
Lam Research	SE	61%	39%	77%	23%	100%	0%	0%	42%	4%	47%	6%	0%	0%	5%	95%
Lattice Semiconductor	SE	33%	67%	15%	85%	100%	0%	0%	59%	0%	38%	3%	0%	0%	8%	92%
Lenovo	CHOE	25%	75%	22%	78%	100%	0%	0%	48%	0%	52%	0%	0%	0%	2%	98%
Lexmark	CHOE	52%	48%	81%	19%	100%	0%	0%	58%	4%	36%	2%	2%	2%	12%	83%
Linear Technology	SE	36%	64%	20%	80%	100%	0%	0%	62%	9%	25%	4%	2%	2%	2%	94%
Logitech international	CHOE	31%	69%	39%	61%	100%	0%	0%	29%	20%	40%	11%	0%	0%	3%	97%
LSI Corp	SE	38%	62%	44%	56%	95%	0%	5%	43%	10%	43%	5%	0%	0%	5%	95%
Lundbeck	PH	34%	66%	47%	53%	89%	7%	4%	22%	48%	11%	19%	41%	0%	41%	19%
Lupin	PH	33%	67%	0%	100%	100%	0%	0%	33%	0%	33%	33%	33%	0%	0%	67%
Macronix International	SE	64%	36%	57%	43%	100%	0%	0%	49%	6%	38%	7%	0%	0%	12%	88%

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
Marvell Technology	SE	28%	72%	0%	100%	100%	0%	0%	50%	0%	46%	4%	0%	0%	13%	88%
Maxim Integrated Products	SE	21%	79%	17%	83%	100%	0%	0%	33%	0%	67%	0%	0%	0%	17%	83%
MediaTek	SE	21%	79%	38%	62%	99%	1%	0%	47%	14%	29%	11%	1%	0%	11%	89%
MediGene	BIO	0%	100%	0%	100%	100%	0%	0%	50%	25%	25%	0%	25%	0%	25%	50%
Medivir	PH	9%	92%	9%	92%	100%	0%	0%	50%	0%	0%	50%	0%	0%	100%	0%
MEMC Electronics Materials	SE	52%	48%	27%	73%	100%	0%	0%	41%	18%	35%	6%	0%	0%	0%	100%
MerckDE	PH	19%	81%	57%	43%	96%	0%	4%	22%	31%	28%	19%	33%	0%	43%	24%
MerckUS	PH	21%	79%	66%	34%	89%	5%	6%	15%	35%	21%	29%	40%	1%	9%	50%
Mercury Computer Systems	SE	0%	100%	0%	100%	100%	0%	0%	0%	0%	100%	0%	0%	0%	0%	100%
Merz	PH	53%	47%	30%	70%	100%	0%	0%	40%	40%	0%	20%	60%	20%	20%	0%
Micrel	SE	12%	88%	12%	88%	100%	0%	0%	65%	6%	29%	0%	0%	6%	6%	88%
Microchip Technology	SE	12%	88%	13%	87%	100%	0%	0%	50%	26%	13%	11%	3%	0%	8%	89%
Micron Technology	SE	44%	56%	80%	20%	100%	0%	0%	41%	2%	53%	3%	0%	0%	4%	96%
Micronic Mydata	SE	76%	24%	75%	25%	93%	0%	7%	50%	0%	43%	7%	7%	0%	7%	86%
Microsemi	SE	54%	46%	0%	100%	100%	0%	0%	67%	0%	33%	0%	0%	0%	0%	100%
Mitel Networks	TCE	0%	100%	0%	100%	100%	0%	0%	43%	0%	57%	0%	0%	0%	29%	71%
MKS Instruments	SE	24%	76%	43%	57%	100%	0%	0%	56%	4%	38%	2%	0%	0%	6%	94%
Morphosys	BIO	75%	25%	75%	25%	0%	0%	100%	0%	0%	100%	0%	0%	100%	0%	0%
Motorola	TCE	1%	99%	88%	12%	99%	0%	1%	42%	6%	45%	7%	2%	0%	7%	92%
Mylan	PH	73%	27%	70%	30%	100%	0%	0%	40%	20%	0%	40%	0%	0%	20%	80%
NCR	CHOE	30%	70%	42%	58%	100%	0%	0%	26%	23%	32%	19%	0%	0%	6%	94%
Nektar	BIO	15%	85%	4%	96%	100%	0%	0%	0%	0%	33%	67%	33%	0%	0%	67%
Neopost	CHOE	39%	61%	27%	73%	100%	0%	0%	28%	0%	61%	11%	0%	0%	17%	83%
NetApp	CHOE	70%	30%	61%	39%	100%	0%	0%	53%	0%	47%	0%	0%	2%	2%	96%

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
NeuroSearch	BIO	78%	22%	78%	22%	100%	0%	0%	0%	67%	0%	33%	100%	0%	0%	0%
NicOx	PH	24%	76%	53%	47%	100%	0%	0%	14%	43%	14%	29%	71%	0%	14%	14%
Nokia	TCE	12%	88%	90%	10%	100%	0%	0%	25%	13%	42%	19%	1%	1%	4%	94%
Novartis	PH	20%	80%	77%	23%	90%	3%	7%	20%	37%	16%	28%	47%	2%	36%	16%
NovoNordisk	PH	40%	60%	80%	20%	95%	0%	5%	27%	9%	59%	5%	0%	0%	23%	77%
NPS	BIO	81%	19%	0%	100%	33%	0%	67%	33%	33%	33%	0%	33%	0%	0%	67%
NVIDIA	SE	18%	82%	35%	65%	100%	0%	0%	50%	1%	44%	5%	0%	0%	4%	96%
OKI Electric	TCE	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%
OmniVision Technologies	SE	66%	34%	68%	32%	100%	0%	0%	50%	12%	24%	14%	0%	0%	14%	86%
ON Semiconductor	SE	53%	47%	33%	67%	99%	0%	1%	48%	2%	47%	3%	0%	2%	19%	80%
Onyx	PH	5%	95%	5%	95%	100%	0%	0%	0%	33%	33%	33%	0%	0%	33%	67%
Orexo	PH	0%	100%	0%	100%	100%	0%	0%	0%	0%	0%	100%	100%	0%	0%	0%
Otsuka	PH	50%	50%	48%	52%	95%	5%	0%	21%	42%	5%	32%	37%	0%	32%	32%
PACE	TCE	33%	67%	33%	67%	100%	0%	0%	0%	0%	100%	0%	100%	0%	0%	0%
Pantech	TCE	0%	100%	0%	100%	100%	0%	0%	0%	100%	0%	0%	0%	0%	100%	0%
Perrigo	PH	17%	83%	17%	83%	100%	0%	0%	100%	0%	0%	0%	17%	0%	33%	50%
Pfizer	PH	31%	69%	80%	20%	97%	1%	2%	23%	34%	19%	25%	49%	0%	12%	39%
Pitney Bowes	CHOE	38%	62%	70%	30%	99%	0%	1%	32%	11%	40%	18%	0%	0%	13%	87%
Plantronics	TCE	49%	51%	23%	77%	100%	0%	0%	15%	8%	38%	38%	0%	0%	0%	100%
PMC-Sierra	SE	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%
Polycom	TCE	60%	40%	48%	52%	100%	0%	0%	31%	0%	54%	15%	0%	0%	0%	100%
Powerwave Technologies	TCE	74%	26%	25%	75%	100%	0%	0%	45%	0%	55%	0%	0%	0%	0%	100%
Qiagen	BIO	63%	37%	77%	23%	89%	11%	0%	67%	0%	28%	6%	17%	0%	22%	61%
Qlogic	SE	13%	88%	0%	100%	100%	0%	0%	25%	0%	75%	0%	0%	0%	0%	100%

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
Qualcomm	TCE	20%	80%	83%	17%	100%	0%	0%	41%	2%	56%	2%	0%	0%	3%	96%
Quantum	CHOE	61%	39%	80%	20%	100%	0%	0%	49%	8%	36%	7%	0%	0%	13%	87%
Radiall	TCE	0%	100%	0%	100%	100%	0%	0%	0%	100%	0%	0%	0%	0%	0%	100%
Rambus	SE	49%	51%	52%	48%	99%	1%	0%	49%	4%	47%	0%	0%	0%	4%	96%
Regeneron	BIO	83%	17%	83%	17%	100%	0%	0%	43%	29%	14%	14%	14%	0%	57%	29%
Research in motion	TCE	30%	70%	52%	48%	100%	0%	0%	49%	2%	47%	2%	64%	0%	9%	26%
RF Micro Devices	SE	32%	68%	13%	87%	89%	0%	11%	11%	0%	78%	11%	0%	0%	0%	100%
Ricoh	CHOE	50%	50%	69%	31%	99%	0%	1%	49%	0%	50%	1%	3%	0%	2%	95%
Roche	PH	14%	86%	82%	18%	97%	2%	1%	41%	14%	38%	8%	35%	0%	21%	44%
Rohm	SE	23%	77%	8%	92%	100%	0%	0%	62%	5%	32%	2%	3%	0%	6%	91%
SanDisk	SE	49%	51%	83%	17%	98%	0%	2%	40%	0%	59%	1%	0%	0%	0%	100%
Seagate Technology	CHOE	47%	53%	63%	37%	100%	0%	0%	55%	5%	33%	6%	4%	10%	5%	82%
Seiko Epson	CHOE	9%	91%	39%	61%	93%	1%	6%	57%	2%	39%	1%	9%	1%	13%	76%
Shire	PH	60%	40%	52%	48%	100%	0%	0%	50%	20%	30%	0%	10%	0%	50%	40%
Sierra Wireless	TCE	28%	72%	19%	81%	100%	0%	0%	31%	38%	15%	15%	0%	0%	8%	92%
Silence	BIO	50%	50%	0%	100%	100%	0%	0%	0%	100%	0%	0%	0%	0%	0%	100%
Silicon Graphics International	CHOE	50%	50%	0%	100%	100%	0%	0%	50%	25%	25%	0%	0%	0%	0%	100%
Silicon Image	SE	39%	61%	20%	80%	89%	0%	11%	56%	11%	33%	0%	11%	11%	0%	78%
Silicon Laboratories	SE	21%	79%	20%	80%	100%	0%	0%	49%	2%	47%	2%	0%	0%	2%	98%
Smartrac	SE	29%	71%	17%	83%	100%	0%	0%	0%	0%	100%	0%	0%	0%	0%	100%
Soitec	SE	69%	31%	15%	85%	86%	10%	3%	48%	3%	48%	0%	31%	3%	14%	52%
Sonus Networks	TCE	0%	100%	0%	100%	100%	0%	0%	50%	50%	0%	0%	0%	0%	0%	100%
Spirent Communications	TCE	23%	77%	0%	100%	100%	0%	0%	50%	0%	50%	0%	10%	0%	10%	80%
Stada	PH	0%	100%	0%	100%	100%	0%	0%	75%	0%	25%	0%	75%	0%	0%	25%

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
Stats ChipPAC	SE	67%	33%	0%	100%	100%	0%	0%	0%	13%	80%	7%	0%	0%	7%	93%
STMicroelectronics	SE	0%	100%	73%	27%	91%	1%	8%	50%	2%	45%	3%	35%	0%	26%	38%
Sun Pharmaceutical Industries	PH	0%	100%	0%	100%	67%	0%	33%	33%	0%	33%	33%	33%	0%	33%	33%
Suss MicroTec	SE	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%
Synaptics	CHOE	72%	28%	66%	34%	100%	0%	0%	63%	0%	25%	13%	0%	0%	0%	100%
Taisho Pharmaceutical	PH	75%	25%	25%	75%	100%	0%	0%	50%	50%	0%	0%	0%	0%	50%	50%
Takeda Pharmaceutical	PH	16%	84%	4%	96%	100%	0%	0%	44%	22%	11%	22%	22%	11%	44%	22%
Telit Communications	TCE	0%	100%	0%	100%	100%	0%	0%	0%	0%	100%	0%	0%	100%	0%	0%
Tellabs	TCE	41%	59%	65%	35%	100%	0%	0%	50%	24%	15%	12%	6%	0%	15%	79%
Teradyne	SE	52%	48%	65%	35%	100%	0%	0%	63%	2%	34%	2%	2%	0%	9%	89%
Tessera Technologies	SE	60%	40%	40%	60%	100%	0%	0%	5%	5%	86%	5%	0%	0%	5%	95%
Teva	PH	44%	56%	44%	56%	95%	0%	5%	34%	20%	46%	0%	22%	2%	10%	66%
Texas Instruments	SE	10%	90%	75%	25%	99%	0%	1%	43%	6%	44%	7%	0%	0%	15%	84%
Theravance	PH	66%	34%	42%	58%	100%	0%	0%	0%	38%	13%	50%	25%	0%	63%	13%
ThromboGenics	BIO	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	0%	0%	100%	0%
Transgene	BIO	40%	60%	60%	40%	0%	100%	0%	0%	100%	0%	0%	0%	0%	100%	0%
Triquint Semiconductor	SE	21%	79%	21%	79%	100%	0%	0%	13%	7%	60%	20%	0%	0%	13%	87%
UCB	PH	61%	39%	26%	74%	100%	0%	0%	17%	17%	33%	33%	17%	0%	50%	33%
United	BIO	0%	100%	0%	100%	100%	0%	0%	50%	0%	50%	0%	0%	0%	0%	100%
Valeant	PH	0%	100%	0%	100%	100%	0%	0%	33%	0%	67%	0%	0%	0%	33%	67%
VeriFone Systems	CHOE	0%	100%	0%	100%	100%	0%	0%	50%	0%	50%	0%	0%	0%	50%	50%
Vertex	BIO	84%	16%	69%	31%	100%	0%	0%	32%	20%	28%	20%	40%	12%	4%	44%
Viasat	TCE	31%	70%	14%	86%	100%	0%	0%	17%	0%	83%	0%	0%	0%	0%	100%
Vitesse Semiconductor	SE	0%	100%	0%	100%	75%	25%	0%	50%	0%	50%	0%	25%	0%	0%	75%

Company	Industry segment	CORE	NONCORE	EXPLOIT	EXPLOR	CLOSED	OPEN_SCI	OPEN_IND	INCR	MOD	ARCH	RAD	LOW	TECH	MARK	HIGH
VTech	TCE	36%	64%	6%	94%	100%	0%	0%	38%	0%	62%	0%	0%	0%	9%	91%
WarnerChilcott	PH	61%	39%	0%	100%	100%	0%	0%	33%	0%	0%	67%	0%	0%	33%	67%
Watson	PH	92%	8%	46%	54%	100%	0%	0%	0%	25%	50%	25%	0%	0%	25%	75%
Western Digital	CHOE	5%	95%	5%	95%	100%	0%	0%	57%	0%	43%	0%	0%	0%	14%	86%
Wolfson Microelectronics	SE	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	0%	0%	0%	100%
Xerox	CHOE	31%	69%	84%	16%	100%	0%	0%	46%	7%	40%	6%	0%	0%	8%	92%
Xilinx	SE	33%	67%	72%	28%	100%	0%	0%	42%	0%	58%	0%	0%	0%	0%	100%
Xyratex	CHOE	0%	100%	0%	100%	100%	0%	0%	100%	0%	0%	0%	100%	0%	0%	0%
Zeltia	PH	56%	44%	40%	61%	75%	13%	13%	13%	38%	13%	38%	25%	0%	75%	0%
ZyXEL Communications	TCE	29%	71%	0%	100%	100%	0%	0%	29%	14%	43%	14%	86%	0%	0%	14%