COME CLOSE AND CO-CREATE

PROXIMITIES IN PHARMACEUTICAL INNOVATION NETWORKS

Come Close and Co-Create

Proximities in Pharmaceutical Innovation Networks

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof.dr. S.W.J. Lamberts

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
6 november 2008 om 16:00 uur

door

Sandra Magdalena Wilhelmina Phlippen

geboren te Kerkrade

ERASMUS UNIVERSITEIT ROTTERDAM

Promotiecommissie

Promotor(en):	**Prof.dr. G.A. van der Knaap**
	Prof.dr. H. Commandeur
Overige leden:	**Dr. H.P.G. Pennings**
	Prof.dr.ing. F.A.J. van den Bosch
	Prof.dr. E. Claassen

Voor Rotterdam

Voorwoord (acknowledgements in Dutch)

6 november 2008. Daar sta ik, Sandra Phlippen, geboren te Kerkrade als dochter van José en Josef Phlippen, zusje van Stephanie Phlippen. Wauw, wat ben ik trots. De mijlpaal van een promotie is zo groot dat het je bij je oorsprong stil doet staan. Hoewel een promotie geen verrassing hoeft te zijn gezien mijn eerdere activiteiten op het gymnasium en de twee afgeronde studies erna, zullen de meeste mensen mij niet direct als academicus inschatten, maar eerder als sociale netwerker. Beide kanten horen bij mij en dat ze niet altijd even gemakkelijk samen gaan komt tot uiting in een duidelijk 2-perioden promotie model.

Tijdens de eerste helft van mijn promotie overheerste mijn sociale kant. Inhoudelijke verkenning, verbreding en tenslotte verdieping van mijn kennis gingen prima samen met aio-clubjes, feestjes en vele andere avonturen. Voor de inspiratie die werd opgedaan tijdens talloze weekendtrips en congressen was meer nodig dan mijn luttele aio-salaris, waarvoor ik pappa dan ook zeer erkentelijk ben. Centrale figuren van het eerste uur die mijn hedonistische levensstijl ondersteunden zijn onder andere Stefano, Rob, Vali, Tibor, Sebi, Viktoria, Stefan, Daina en Kevin. Met Francesco Ravazzolo, mijn kamergenoot, heb ik lief en leed gedeeld. Met name door hem was onze kamer H16-10 lange tijd het sociale epicentrum van het Tinbergen Instituut. Een klein aantal aio's zijn mij dierbaar. Streepje, Michiel, Francesco, Elaine, Gus, Bas, Flor en Romy weten dat. Met Ward en Jan Frederik heb ik gouden koffiekamertijden gedeeld, waarbij mijn links-liberale politieke standpunt alleen maar sterker is geworden. Vrouwen zijn zeldzaam bij het Tinbergen Instituut, maar Silvia en Mariëlle zijn twee briljantjes.

In het najaar van 2004 ontstond de verdieping in mijn werk tijdens een werkbezoek aan de universiteit van Florence en in Lucca. Massimo Riccaboni heeft mij geïnspireerd en getraind in het netwerk denken. Sindsdien kijk ik gefascineerd door een netwerkbril naar de wereld. Andere brildragers zijn Marco van der Leij, Skye Bender-de-Mol, Andrea Knecht, Marielle Non, Martijn Burger, Bas Karreman, Ronald Wall en mijn promotor Bert van der Knaap. Mijn laatste dankwoorden voor

deze eerste periode zijn voor de TI staf. Maarten, Ine, Dave, de 2 Cariene's, Jaap en Arianne leidden een mooi instituut.

Toen kwam Robert Dur. Zijn liefde bracht rust en concentratie die nodig waren voor het werkelijke schrijven van een proefschrift. Maar nog belangrijker dan dat, Robert maakte al het andere relatief en hielp mij om twijfel om te zetten in vastberadenheid. Papers werden geschreven, congressen bezocht en positie genomen in het wetenschappelijke debat. Dit alles, en het college geven aan 800 eerstejaars economie studenten heeft mij gevormd tot wie ik nu ben. Mijn promotoren Bert van der Knaap en Harry Commandeur hebben hierin een belangrijke rol gespeeld. Bert gaf mij de ruimte om te groeien, hielp mij op mijn werk reflecteren en samen kregen we ideeën waar we enthousiast van werden. Ik heb je kamer zowel huilend als juichend verlaten, hoewel je dat eerste hoop ik nooit hebt gezien. Gelukkig was er na je kamer altijd Nita.. Ik ben je dankbaar voor alles Bert. Harry Commandeur is een fantastische tweede promotor. Alles wat ik schreef werd grondig gelezen en we hebben vele uren zitten bomen over interpretaties, ideeën en uitbreidingen. Mijn werk is beter geworden door zijn commentaar.

Een hoogtepunt was mijn onderzoeksproject bij McKinsey & Company in Brussel en Louvain-la-Neuve. Van An Vermeersch en Daniel Berhin heb ik bijvoorbeeld geleerd dat het bottom-up denken in de wetenschap veel sneller en efficiënter kan. Amanda, Alain, Filip, Arlene, Natu en de anderen in de pharma practice waren fijne collega's. Daarbuiten ben ik Bernard Paque, Leigh Weiss en Olivier van Rietpaap dankbaar voor hun bereidheid tot kennisuitwisseling en introducties.

In beide perioden van mijn promotie en al ver daarvoor waren er mijn hartsvriendinnen die onvoorwaardelijk al mijn uitspattingen en onhebbelijkheden accepteren. Amber, Etje, Francesca, Johanna, Marloeschka, en Streep, wat fijn dat jullie er zijn. Als laatste wil ik de Commissieleden hartelijk bedanken voor hun inzet en het Erasmus Trustfonds voor de royale financiële bijdragen aan mijn congres- en werkbezoeken in het buitenland.

En toen was er Elodie. Ook al heb je mijn promotie aanzienlijk vertraagd, in je moeder heb je je grootste fan gevonden. Hoewel het steeds vechten blijft met je vader om die positie. Over 2 weekjes ben je een jaar in ons leven. Het mooiste jaar.

Zittend in Dudok Rotterdam schrijf ik dit voorwoord. Rotterdam is de stad waarin dit alles heeft plaatsgevonden. Rotterdam is mijn stad. De ruimte, de wind, de rauwheid, de stilte en de stedelijkheid brengen mij troost en inspireren. Het is de stad van sterke, individuele koppen aan wie ik dit proefschrift opdraag.

Sandra Phlippen

Rotterdam, 2 september 2008

Table of contents

Chapter 1   Introduction

1.1 Motivation, aim, and scope

In studying firm behavior, economists tend to have an under-socialized view of the firm, while sociologists tend to have an over-socialized view of the firm. Socialization in this respect refers to the extent to which a firm is embedded in- and affected by its relational environment. On the one extreme, economists building on transaction costs economics assume the market to be an anonymous environment where firms can behave opportunistically without repercussions. On the other extreme, sociologists argue that the market consists of a dense web of existing and previous transactions in which a firm is embedded, and that this embeddedness is defining the range of strategies a firm can pursue or even perceive[1]. While the former view mainly considers the impact of behavior on a static environment and the latter focuses on the impact of environment on firm behavior, this thesis provides a co-evolutionary perspective, wherein firm behavior and environmental structures affect each other mutually. This dynamic perspective on innovation strategies of firms allows us to take into accounts both ongoing and previous transactions, which together constitutes a firm's relational environment. We study how firm behavior is affected by its network position and how network structures change as a result of firm behavior. We shall focus on the biopharmaceutical industry. Besides the network of collaborations, the environment in which a typical biopharmaceutical firm operates consists of a number of other elements as depicted in figure 1.1. This figure shows the main environmental elements to which a firm in the pharmaceutical industry is subject to.

---

[1] The notion that firms face different 'menus' of technological choices because of the path dependent development trajectories is put forward by Nelson & Winter (1982).

Fig. 1.1 - The environment of the firm
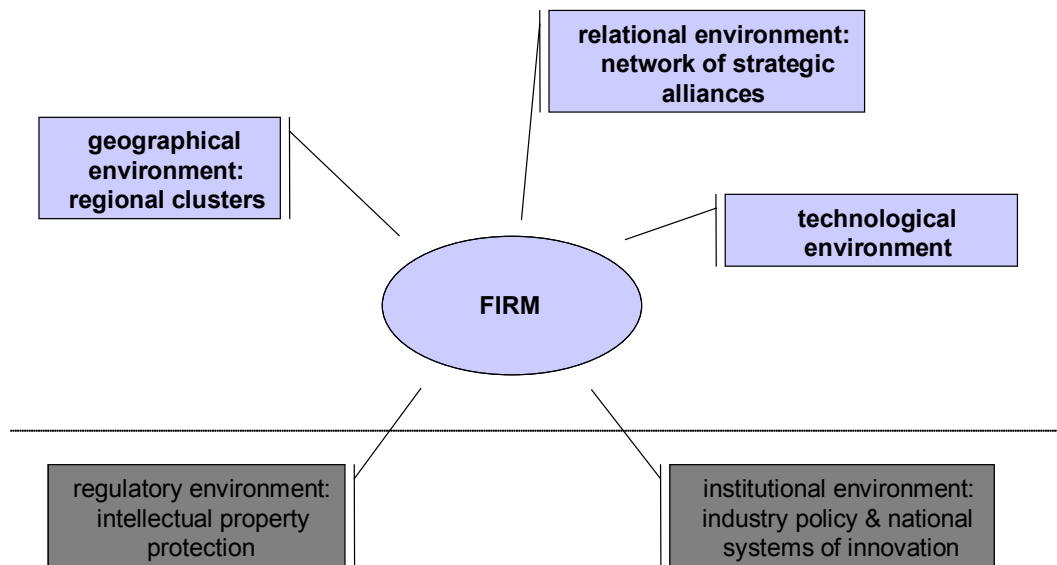
One very important element to the firm is the technological environment, since in the pharmaceutical industry firms are active in a highly volatile technological environment, which requires them to continuously adapt and to respond to new technological developments. The institutional and regulatory environments in which the firm operates are also of great importance. While the institutional and the regulatory environment are evolving rather slowly over time, and are outside the scope of this thesis, they are of great importance for understanding the innovation strategy of the firm. Especially, the issue of intellectual property right protection is of major importance in providing firms with the freedom to manoeuvre outside their boundaries. However, the complexity of this latter issue deserves a dissertation of its own. Finally, there is the geographical environment that we consider in understanding innovation strategies of pharmaceutical firms. In science-based industries, such as the pharmaceutical industry, the physical environment of the firm is particularly important to understand who benefits from (mainly) public (scientific) knowledge spillovers.

Within these environments, firms face the decision to either internalize their production, to buy from the market, or to ally with partners. For the development of new drugs, which is a lengthy, costly and highly uncertain process, firms develop

strategic portfolios in which they make, buy *and* ally simultaneously. Especially the large, established pharmaceutical firms, which dominate the industry have for the last decade been searching for new strategic approaches to combine in-house activities with boundary spanning activities[2] [3]. The main reason being the continuous decrease of new drugs in their pipelines combined with expiring patent protection on blockbuster drugs. For other organizations in the industry, such as smaller biotech firms or academic organizations, the willingness of pharmaceutical firms to engage in boundary spanning activities provides interesting opportunities for innovation. Boundary spanning activities between organizations in the pharmaceutical industry can be depicted as a network of organizations and their inter-firm relations. Relations in this network represent strategies of organizations to access knowledge externally. This dissertation focuses on how geographical-, relational, and cognitive (technological) proximity between two organizations affects a firm's access to external knowledge and hencefort their innovation strategies.

In this introductory chapter we first provide a short overview of the theoretical discussion on firm strategy and the impact of proximities. Next, we introduce the reader to our empirical setting by explaining the production process of drug development and by providing a short historical overview. Finally the outline of the remainder of this book is presented to the reader.

1.2 Innovation strategy of the firm

1.2.1 Dynamic capabilities

Firms operating in environments of rapid technological change are required to develop dynamic capabilities in order to survive the waves of creative destruction in an industry (Teece, et al., 1997, Prahalad & Hamel, 1990; Penrose, 1995; Schumpeter, 1942). Dynamic capabilities refer to a firm's ability to exploit its existing internal

---

[2] For example Glaxo Smith-Kline has set up various centers of excellence for external drug discovery where specific attempts are made to align in-house R&D with external drug discovery efforts. *Financial Times,* June 8 2008.
[3] Elli Lilly, to give an example of one of the largest drug developing firms, is one of those firms that are continuously searching for new strategies to boost their drug discovery output. In 2001 it has spun out a 'virtual' firm named InnoCentive, which poses drug discovery problems on the web including fees for 'solvers' (Travis, 2008).

knowledge and simultaneously explore new (external) knowledge to address changing environments (Teece et al, 1997; O'Reilly & Tushman, 2004). Together exploration and exploitation can be seen as a cycle of innovation, where new products, new technologies or new applications are constantly being *explored* and *exploited* in new markets (Nooteboom, 2000). Ideas, institutions, and organizational routines first settle down in a 'best practice' or a dominant design. That shows up the limits of validity as well as indications for novel combinations, which break down existing structures. This leads to a next round of convergence into a dominant design.

By analyzing a firm's innovation strategy in terms of dynamic capabilities we encompass several theoretical views of the firm. The transaction costs theory of the firm and the resourced based view of the firm provide insights on how firms choose an optimal governance structure that minimizes the costs and maximizes the value of each individual transaction. To understand the potential complementarities between individual transactions in a firm's portfolio (or a network) we further consider 3 other theoretical perspectives on firm strategy: the relational view of the firm, which emphasizes the whole portfolio of transactions as a unique resource of the firm. The embeddedness view of the firm, which focuses on the costs minimizing features of the network surrounding the firm. Finally, we consider the real options approach to firm strategy, which adds additional arguments to how a firm can strategize on its portfolio. The real options approach mainly argues that firms can capitalize on transaction value and simultaneously minimize transaction costs by confronting uncertainty with flexibility, rather than avoiding uncertainty (Kogut, 1991; Leiblein, 2003)

1.2.2 Make, buy or ally

The strategy of a firm on how it is going to govern its transactions is one of the core issues in economics. Possible governance structures range from full integration, which means that transactions are performed internally to arms-length transactions in the market, which refers to a completely disintegrated structure. In between these two extremes are long term contracts, strategic alliances & joint ventures, and parent subsidiary relationships ranging from less- to more integrated governance structures

(Besanko, 2007). Figure 1.2 plots the range of possible transactions ranging from full integration to pure market transactions.

Fig. 1.2 - Range of mechanisms to govern transactions
Based on Yoshino & Rangan (1995)

Strategic alliances

Contractual agreements   Equity arrangements

Traditional   Non-traditional   Creation of entity   Dissolution of entity

**Technology
acquisition**   **Collaborative R&D**

                       Co-development   **Joint venture**   **Merger &
                                                                               acquisition**

Licensing
                       **Joint marketing &
                       distribution**

Inter-firm linkages
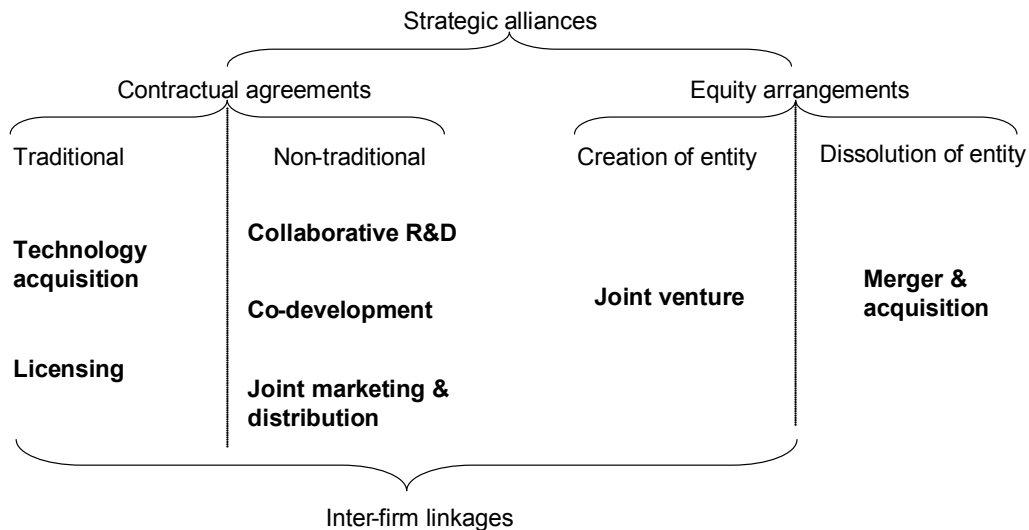
By and large, these governance structures vary along three dimensions: efficiency (transaction value), costs (transaction costs), and the frequency of transactions.

Efficiency

According to transaction costs economics (TCE), markets are generally more efficient than hierarchies, since in a market a product can serve multiple customers and achieve scale and scope economies. On the other hand the firm (representing a hierarchy) is argued to have more superior coordinative attributes and information processing abilities (Gulati & Singh, 1998). When transactions are geared toward innovation, coordination and information processing becomes increasingly important. This is because transactions aimed at innovation contain rather large amounts of tacit (embedded) knowledge, which is not perfectly tradable and requires more coordination and processing capabilities (Kogut & Zander, 1992). While performing a transaction internally might be more efficient in terms of coordinating and processing tacit knowledge, it might be impossible to acquire these resources within the firm

given the difficulties of mobilizing them. Indeed, the resource based view (RBV) of the firm argues that firms exist because they possess valuable, rare, hard to imitate and immobile resources such as tacit (embedded) knowledge (Barney, 1991). Because these resources are sticky, firms are heterogeneous, which in turn enables rent creation from transactions outside the boundaries of the firm. This latter requirement is particularly prevailing when firms are active in technologically volatile markets. When transactions are focused on innovation, stickiness prevents a firm to fully integrate resources internally, but it also prevents the transfer of these resources in pure market transactions. In other words, when transactions entail rather large amounts of tacit (embedded) knowledge (which is sticky) strategic alliances between firms might be most efficient. With this intermediate form of governance a firm can still achieve scale and scope economies by interacting with multiple partners and at the same time allow sticky information to gradually be transferred between alliance partners. Finally, Dyer and Singh (1998) introduce the notion of how a complete portfolio of boundary spanning transactions can be viewed as a unique resource of the firm. Taken together, the portfolio of transactions (or the ego-network) can provide a firm with relational rents, which creates a competitive advantage in an industry. Relational rents are defined as "supernormal profit jointly generated in an exchange relationship that cannot be generated by either firm in isolation and can only be created through the joint idiosyncratic contributions of the specific alliance partners" (Dyer & Singh, 1998, p.662). While the relational view mainly emphasizes value creation as a competitive resource of the firm, in the following section we will emphasize how the portfolio of boundary spanning transactions prevents knowledge appropriation and thereby enables a *sustainable* competitive advantage.

Costs

There are two main economic hazards involved in transactions. These are the danger of opportunistic behavior by transaction partners, and the danger of being unable to overview the whole transaction process, which is referred to as bounded rationality. Opportunistic behavior can entail the unintended appropriation of knowledge (knowledge leakage) or the unwillingness to pay once a transaction has started. Governance structures vary in the way in which, and the extent to which they mitigate

the dangers of a transaction as described above. According to transactions costs economics (Williamson, 1975), market transactions can in principle mitigate opportunistic behavior through contracts that specify the mutual agreements between agents in a market. However, given that agents cannot oversee the complete transaction, contracts are by default incomplete. Moreover, when transactions are aimed at exploring novel information, they are often highly specific and their outcome uncertain or hard to determine in advance. In these cases, contracts are even more incomplete and this increases the danger of knowledge leakage. In addition, the uncertainty about the outcome of a transaction ex-ante requires renegotiation along the way, which given the specificity of knowledge, greatly increases the danger of hold-up by the market agent (Leiblein & Miller, 2003). These circumstances lead transaction costs economists to conclude that the transaction should be vertically integrated inside the firm, where "fiat rules the day" (Leiblein, 2003, p. 941). However, real options theory provides compelling arguments for refraining from vertical integration since this could become costly in the future. More specifically, real options theory argues that in environments where technologies change rapidly, vertical integration decreases the flexibility to abandon assets that turn out less valuable after technological change (Pindyk, 1991). If a firm governs its transactions through a portfolio of strategic alliances with various partners it can preserve its flexibility by opting to withdraw from a less promising alliance or it can obtain scope economies by re-applying similar technologies in different relational transactions and applications. This latter strategy has also been recognized as hedging against the dangers of knowledge appropriation because firms create unique 'bundles' of relation specific knowledge, which is hard to imitate[4] (Dyer & Singh, 1998).

Thus far we have seen that TCE considers contracting or, if impossible, internalization to be the answer to dealing with the risks of opportunistic behavior. By analyzing each individual transaction, TCE takes a dyadic view on the innovation strategy of the firm. The real options theory provides a broader perspective by analyzing the whole portfolio of transactions in which a firm is engaged. It argues that uncertainties from the environment can be mitigated by diversifying transactions in such a manner that the firm can flexibly switch between transactions if environmental

---

[4] In the relational view of the firm this is referred to as inter-organizational asset interconnectedness (Dyer & Singh, 1998)

conditions change. The network perspective takes an even broader perspective by not only taking into account the transaction partners of a firm, but also the *partners'* partners of the firm and *their* transactions. The theory argues that the more 'embedded' a firm is in a 'web' of current, previous, and partnering transactions, the less it will have to spend on mitigating the risks of opportunistic behavior (Granovetter, 1973). More precisely, opportunistic behavior is prevented because of reputation information circling around in the network of transactions. Additionally, reputation information enables the enforcement of social norms and eventually 'trust' between transacting partners, which are recognized to be good alternatives for costly contracting.

Transaction frequency

In general one can say that transactions occur more frequently as the firm moves from more integrated (in-house) governance structures to less integrated (market) transactions. Furthermore, contracts are an important part of a transaction regardless of the type of governance structure. However, the more frequent transactions are likely to occur, the more difficult it becomes to make contracts complete (given bounded rationality and increased uncertainty) and the more important alternative approaches to control transaction outcomes become. Within the boundaries of the firm, contracts are enforced and complemented through power differences and fiat. Beyond the boundaries of the firm, where we distinguish arms-length transactions and embedded transactions (strategic alliances), it has been found that in more embedded transactions, contracts are complemented by trust (Uzzi, 1997). Trust builds on heuristics of experience, which accumulates over time (Uzzi, 1997, p.43-44). Finally, by transacting more frequently, partners involved in joint problem solving are better able to overcome asymmetric information between them. The reason for this is that transactions are reciprocated over time, which leads to relation-specific investments by both partners. Although the specificity of investments can easily lead to situations of small-numbers bargaining and subsequent hold-up problems, the situation is unlikely to generate opportunistic behavior by any partner because of mutual dependencies.

Thus far we have discussed the considerations of firms in choosing the appropriate governance structure to encompass their innovation activities. When firms operate in technologically volatile environments, there are convincing arguments for why firms should complement their existing in-house activities with explorative search beyond their boundaries. The question that has remained largely under-exposed is how firms should go about in achieving complementarities between the different transactions they engage in. Put differently, when are boundary spanning transactions usable for achieving innovation performance internally?

1.2.3 Availability and usability of knowledge

In order to effectively use external sources of knowledge, firms first need to gain access to knowledge (availability) and second they need to be able to use or absorb this knowledge (usability). The issue of availability and usability of knowledge has been analyzed by Simon (1958) and was later expanded by Cohen and Levinthal (1989). The strategies that firms develop in their search endeavor for available and usable knowledge has mainly been categorized in terms of local-, and non-local search (Simon, 1958; Nelson & Winter, 1982; Katila & Ahuja, 2002). Local and non-local search represent the ends of a number of continuums that we will call proximities.

Access to / availability of knowledge

The strategies of organizations to acquire access to new knowledge and technologies depend on the type and source of knowledge and on the boundary spanning capabilities of the organization. The source of knowledge can be public or private, and the type of knowledge can be codified or rather tacit. Public, codified knowledge, which is non-excludable and non-rival is found to spill-over to organizations located within geographical proximity of the source of knowledge. Access to private knowledge can be obtained through vertical integration such as labor mobility or mergers & acquisitions or through boundary spanning activities such as strategic alliances or arms-length market transactions. This thesis mainly focuses on strategic alliances between organizations as an important means to obtain access to external

knowledge. Together these strategic alliances form a large network, covering all phases in the production process. While the network of strategic alliances is often claimed to be the locus of innovation, the extent to which new knowledge is actually generated through the alliance depends on the type of strategic collaboration. R&D collaborations or co-development alliances are for example most often alliances whereby all partners involved contribute some of their knowledge to the alliance. When transactions are licensing agreements, knowledge has most likely been created within the principal firm, as a partnering firm pays a license fee to acquire knowledge from the principal firm.

Usability of knowledge / absorption capacity

While access to knowledge is a necessary condition for innovation, it is not a sufficient one. Once a firm has obtained access to either a public or a private source of external knowledge, this knowledge needs to be usable within the organization. According to Cohen and Levinthal (1989), firms need to build a capacity to absorb external information. A firm's absorption capacity is defined as its 'relevant stock of prior knowledge'. 'Prior knowledge' refers to the knowledge that a firm has built internally before tapping into the external source of knowledge. A 'relevant' stock of knowledge can be interpreted as knowledge that is similar to the external knowledge that the firm is aiming to obtain.

1.2.4 Proximities

There are various forms of proximity, which affect an organization's ability to access external knowledge and its ability to absorb this knowledge and turn it into innovation. While it is often stated otherwise, we argue that proximities are important for not only the transfer and absorption of tacit knowledge, but also for codified knowledge. Even when knowledge is codified, the interpretation and internalization of knowledge still requires tacit knowledge and therewith, proximity (Nonaka & Takeuchi, 1995 pp 72; Howells, 2002; Boschma, 2005). The most straightforward form of proximity is geographical proximity, which is the physical distance between two or more actors (either public of private). Cognitive proximity is defined as the

degree of commonality in knowledge domain between two actors. This knowledge domain acts as a framework by which we perceive, interpret, understand and evaluate new information. The cognitive framework is constantly reshaped in the process of interaction with others (Mead, 1934). In practice, cognitive proximity is often measured as technological proximity, which refers to the degree of overlap in the technologies two actors are specialized in.

A third form of proximity is relational or network proximity. Relational proximity refers to how close two actors are in a network. Usually relational proximity is expressed as negative relational distance. Relational distance is a network measure. It starts with the notion that all organizations are nodes and each exchange of knowledge between these organizations is an edge between nodes. The relational distance between any pair of nodes is the number of edges one has to surpass to reach the other. At a distance of 1, two firms are partners, and a distance of 2 implies that two firms share a common third partner. Relational proximity comes very close to the notion of social proximity, which is defined as "a socially embedded relation between agents at the micro-level. Relations between actors are socially embedded when they involve trust based on friendship, kinship and experience" (Boschma, 2005 pp.66). The difference between social and relational proximity is that social proximity describes a dyadic feature of a relation, while relational proximity takes the whole network of previous collaborations with all partners into account. Table 1.1 summarizes the effects of three forms of proximities on the availability and usability of external knowledge.

Table 1.1   – Effect of proximities on availability and usability of
                external knowledge

	Availability of knowledge (access)	Usability of knowledge (absorption capacity)
Geographical proximity	Public (codified) knowledge spillovers are geographically bounded (Jaffe & Trajtenberg 1993)	Faceto-face contacts facilitate the transfer of highly specific tacit knowledge (Zucker, 1996)
Relational (social) proximity	Public knowledge spillovers require relational proximity (and geographical proximity enables social network formation) (Breschi & Lissoni, 2001) Relational proximity creates transitivity and social control which, facilitates new partner formation (Gulati, 1995)	Previous partner experience improves alliance management and increases knowledge transfer within the alliance (Gulati, 1995) Socially embedded relations based on trust facilitate tacit knowledge transfer (Maskell & Malmberg, 1999)
Cognitive proximity	Cognitive proximity is required to access epistemic communities with highly specified tacit knowledge.	Cognitive proximity increases a firms' absorption capacity (Cohen & Levinthal, 1989)

Geographical proximity

Patents, which are highly codified sources of public knowledge, are more likely to build on patents that were applied within geographical vicinity then on patents that are applied at distant locations (Jaffe et al, 1993). This finding is interpreted as evidence of geographically mediated public knowledge spillovers. It implies that firms who aim to access this knowledge benefit from being co-located to the source of knowledge. Geographical proximity has proven to be important for the transfer and absorption of tacit knowledge between organizations in our empirical setting; the biopharmaceutical industry. One obvious reason for this is that the natural excludability of knowledge in this industry requires scientists who have invented a new drug or a technology are needed for further development and exploitation of the invention. The dual occupation of these scientists creates human capital immobility and thereby induces knowledge transfer within geographical proximity. Co-located organizations are more likely to engage in long-term face-to-face contacts. These contacts create mutual trust, shared social norms, and an epistemic community that are important for successful transfer and absorption of tacit knowledge (Zucker, 1996).

Relational proximity

While Jaffe et al (1993) found that patents are more likely to be cited by inventors that are geographically proximate, it seems, that relational proximity between organizations is even a better predictor of who cites (and accesses knowledge of) whom (Breschi & Lissoni, 2001). It is generally acknowledged that geographical proximity alone is not sufficient for firms to enter knowledge transfer arrangements or to benefit from spillovers, it is most likely that co-location strengthens other dimensions of proximity (Zucker, 1996). For instance, a minimum amount of cognitive proximity is deemed necessary for absorption of external knowledge, and co-location is found to induce local processes of imitation and selection, which create cognitive proximity (Boschma, 2004).

At the level of the firm, Gulati argues that relational proximity facilitates the formation of new collaborations between organizations. Especially when a firm's network of collaborations has a clique-like structure (i.e. its existing partners collaborate as well), reputation information circulates and functions as a social control mechanism against opportunistic behavior. Social control in turn helps firms engage in risky undertakings such as the start of a collaboration with a new 'unknown' partner. Once collaboration has started, the actors involved may become more 'socially close' as the time spent collaborating increases. This includes for example the formation of direct communication lines between technical staff of two organizations or the development of shared jargon. As a result, externally sourced knowledge can be more fine-tuned to the needs of the organization, which increases its ability to absorb external knowledge and ultimately to innovate.

Cognitive proximity

Cognitive proximity is found to positively affect an organization's ability to access and use external information. Within the organizational learning literature it has been shown that new knowledge acquisition is more successful when internal capabilities are similar to new knowledge that is being sought externally (Dussauge & Garrette,

1999). Especially when knowledge is highly complex and tacit, specialized epistemic communities evolve, which require a high level of cognitive proximity to participate. In science, for example, epistemic communities evolve around specific sub-disciplines within a scientific field. Obtaining access to such communities through e.g. conference participation requires a very high level of knowledge similarity.

Too much proximity and interaction effects

While emphasis has mainly been on the positive effects of proximity, it is undisputable that too much proximity can have detrimental effects for innovation performance or even the likelihood of obtaining access to external knowledge. The negative effects of too much proximity become apparent when the interaction effects between different forms of proximity are taken into account. For example, firms in dense clusters of co-located organizations can suffer from lock-inn effects due to redundancy or from unintended knowledge flows *if* these organizations are densely connected through more or less formal relations. In this case, geographical proximity coincides with relational proximity. Additionally, the probability of lock-inn effects or unintended knowledge flows is even higher when organizations are also active in a similar knowledge domain, an argument, which adds cognitive proximity to the equation.

Furthermore, the usability of external knowledge increases with the degree of cognitive proximity up to a certain point. A study by Wuyts (2005) reveals the inverted u-shaped relation between cognitive proximity between two firms and their ability to learn from each other. Indeed, knowledge similarity initially increases a firm's ability to absorb external knowledge, but too much overlap in two actors' knowledge domain reduces the newness of information and is found to negatively affect the ability to be innovative.

At the level of a cluster there is empirical ambiguity about whether specialization (representing cognitive proximity between clustered organizations), or diversity amongst co-located firms induces knowledge transfer and innovation performance. One the one hand there are studies showing the benefits of flexible specialization (Amin & Thrift, 1992; Piore & Sabel, 1984). Flexible specialization means that firms and other actors in a cluster specialize in a product line, but have the ability to shift to

other related lines with similar technologies rather instantaneously. As a result, these clusters are rather homogeneous in terms of technologies, while complementary knowledge is exchanged between firms inside the cluster intensively. On the other hand, there are studies promoting technological diversity (Prevezer, 1997; Cooke, 2001). Clusters in which firms are technologically diversified appear to encourage cross-fertilization of ideas and technologies between firms, which leads to more innovation and firm growth.

Finally there is contradictory theoretical and empirical evidence about the effect of relational proximity on a firms' innovation capacity. In these studies relational proximity and cognitive proximity have not been clearly separated and are mainly referred to as local-, and non-local search. Local search implies that firms search for new knowledge within the network of existing and previous collaborations. According to Cowan (2005) local search enables firms to access knowledge that it can integrate with its existing capabilities. In a local network, so the author argues, shared language, norms, and collaboration processes have developed, which are particularly valuable when new and tacit knowledge is to be exchanged. Non-local search through 'distant' linkages in a network are considered of major importance for innovation by Sidhu, Commandeur and Volberda (2007). These authors provide evidence that distant linkages provide a firm with access to more novel information, which is crucial for innovation.

1.3 Empirical setting: the biopharmaceutical industry

The biopharmaceutical industry is one of the most interesting industries to study firm innovation. The main reason for this is that innovations occur frequently and even radical innovations occur regularly. For the production of a new drug there are several different types of organizations involved, such as biotech firms, pharmaceutical organizations, academic institutes and public-, and private financial organizations. While a large part of the production process is still being carried out within the boundaries of the firm, there is an abundance of empirical evidence that the importance of innovation taking place outside the boundaries of the firm is increasing (Hagedoorn, 2001; Orsenigo et al, 2001; Phlippen & Riccaboni, 2008). Boundary spanning innovation activity is of interest to us for two reasons. First, it allows us to

monitor innovations and their antecedents more closely. In-house R&D is often subject to considerable scrutiny before made public. Second, inter-organizational innovation activities allow us to study industrial players and their activities as a network of innovators, which provides valuable complementary insights to standard (actor-based) economic thinking about industrial organization. Furthermore, amongst the strategic alliances that make up our network, the two main processes of innovation, being explorative alliances and exploitative alliances, are clearly distinguishable (March, 1991; Jansen, van den Bosch & Volberda, forthcoming). Finally, because of the science driven features of new drug development, a large part of the information that is exchanged between organizations is tacit. Tacit information transfer is, as we have discussed in the previous section, assumed to be strongly affected by different forms of proximity.

1.3.1 Drug development process

The biopharmaceutical industry is strongly innovation driven, as rapid and radical technological innovations threaten to render existing products obsolete within a relatively short time. In order to compete, firms in the biopharmaceutical industry need to continuously develop new or improved drugs and technologies that are valuable and patentable. The skills required for new drug development range from basic research and discovery to clinical testing procedures, manufacturing, marketing and distribution and knowledge of and experience with the regulatory process. Basic research and discovery activities are aimed at exploring new chemical targets or molecular compounds. These exploration activities are strongly science-driven and are often characterized by so-called upstream collaborations between small biotech firms and academia. Once a new compound has been discovered, the clinical testing procedures range from toxicity testing on animals to large scale testing on patients in a controlled hospital environment. At various stages in the process of clinical testing, approval from the Federal Drug Administration is required in order to continue the testing procedures. Because these clinical testing procedures are extremely expensive, large pharmaceutical organizations have specialized in these so-called downstream drug development activities. Once a new drug has received final approval to enter the

market, large pharmaceutical firms exploit the new drug in the market using their well entrenched marketing and distribution channels.

1.3.2 Historical overview

Traditionally, the whole process of drug development, both the exploration of new compounds and the exploitation of existing compounds, has been conducted in-house by large established pharmaceutical firms. The research laboratories of large pharmaceutical firms attracted the best chemists from academia worldwide and deeply specialized knowledge of disease areas was cultivated in these laboratories. From the seventies, a number of radical technological innovations such as the molecular biology revolution and the genomics revolution altered the process of drug development and the role that various organizations play in drug development. In what follows, we will discuss the nature of these innovations, the effect that they had on the drug development process and finally the strategies of organizations to cope with the innovations.

Molecular biology revolution

In the 70s and 80s, drug development was characterized by advancement in chemistry, pharmacology, microbiology and biochemistry, which together led to what was called the molecular biology revolution. Molecular biology potentially enables researchers to understand disease processes at the molecular (genetic) level and to determine the optimal molecular targets for drug intervention (Drews, 2000). New biotechnology drugs based on molecular biology were mainly recombinant proteins and monoclonal antibodies.

Advances in molecular biology originated from American universities and research centers. The first firms exploring biotechnologies were science based dedicated biotechnology firms who were located closely to the academic sources of knowledge in the US in order to collaborate with academic researchers or because of dual occupations. Large pharmaceutical firms, who were at that time still only developing chemical based compounds, realized the potential of the new biotechnologies, and started collaborating with dedicated biotechnology firms. As a result, drug

development was no longer the sole territory of large pharmaceuticals, but was divided into an explorative part and exploitative part. Dedicated biotech firms (DBFs) and academia were collaborating to explore new biotechnology drugs while pharmaceutical firms were collaborating with DBFs to access and subsequently exploit biotech based drugs on the market. European pharmaceutical firms who were eager to access the new biotechnology based drugs were required to collaborate with DBFs in the United States. A number of reasons have been identified for the near to absence of dedicated biotechnology firms in Europe during the 70s and 80s (Owen-smith et al., 2002): first, institutional barriers prevented scientists to become entrepreneurs as was happening in the United States. Second, while private funding through venture capital was not available, public funding from the government was too decentralized. While the National Institute of Health was funding centers of excellence in the US, European funding came from the national level, which could not lead to a critical mass of knowledge. A final reason for the absence of biotech firms in Europe is related to regional knowledge trajectories that were evolving in Europe. Although, US firms involved in biotechnologies were spatially concentrated in a few areas, the innovation activities within these areas were highly diversified. In contrast, the spatial concentration of European activities developed alongside specialized knowledge trajectories, which impeded cross-fertilization of knowledge (E.g. Max Planck institutes).

In the 70s and 80s, access to knowledge had become an indispensable resource for both pharmaceutical firms and for dedicated biotechnology firms. Dedicated biotech firms who were aiming to access academic 'public' knowledge, located near academic centers of excellence in the US to benefit from localized knowledge spillovers. As a result, explorative activities in drug development were regionally concentrated in a few areas. Large pharmaceutical firms who aimed to exploit biotechnology based drugs on the market, formed strategic collaborations with DBFs to access 'private' knowledge they did not possess in-house themselves. While a large part of the drugs that pharmaceutical firms brought to the market still originated in-house, the amount of externally sourced drugs were steadily increasing. From the 70s and 80s, a large network of collaboration activities started to take form in the biopharmaceutical industry. Although European pharmaceutical firms were actively involved in these collaboration activities, most firms had American partners, which prevented the

formation of an intra-European network (Senker, 2004). In 1978, the Single European Act was formed to stimulate intra European research collaboration. This act later developed into the European framework programs.

Genomics revolution

In the beginning of the nineties the Human Genome Project was initiated by the US Government to identify the genes that make up the human DNA and store the genetic information into huge publicly available databases. While the genetic information was publicly available, it was hardly useful since the information came in huge unordered amounts. Scientists working on the decoding of the genome, started developing technologies to store, organize, screen, and subsequently use the genetic information for medical applications. These technologies, referred to as general purpose technologies have fundamentally altered the approach to drug development. Before the genomics revolution, drug research used to take a more qualitative approach, where many scientists worked on developing a few molecules that could effectively interfere with a given target. The genomics revolution suddenly enabled the identification of hundreds of genetic targets and the screening to thousands of molecular compounds simultaneously (Pammoli & Riccaboni, 2002). While this new, more quantitative approach to drug research has not delivered on its promises yet, in the beginning of the nineties pharmaceutical firms were eager to form strategic alliances with general purpose based firms to access these technologies. Today, general purpose based technologies have become fully integrated into the process of drug development. Also, an increasing amount of new drugs on the market are based on biotechnologies.

What we have seen in this historical overview is that since the beginning of the 70s a division of labor took place in the process of drug development. Exploration of new molecular compounds was increasinlgy carried out by academia and dedicated biotech firms. The testing and exploitation in the market has remained the domain of large pharmaceutical firms. The division of labor has made access to knowledge an important prerequisite for firms to innovate. This is especially true during periods when radical innovations occur that affect the industry in fundamental ways. In figure

1.3 we show the relative increase in strategic collaborations in pharmaceutical R&D during the molecular biology revolution and during the genomics revolution.

Fig. 1.3 ‑ Percentage change in R&D collaborative   agreements
         (E ‑ Edges) and organizations (V ‑ Vertices) in the
         biopharmaceutical R&D network (1976 ‑ 2002)
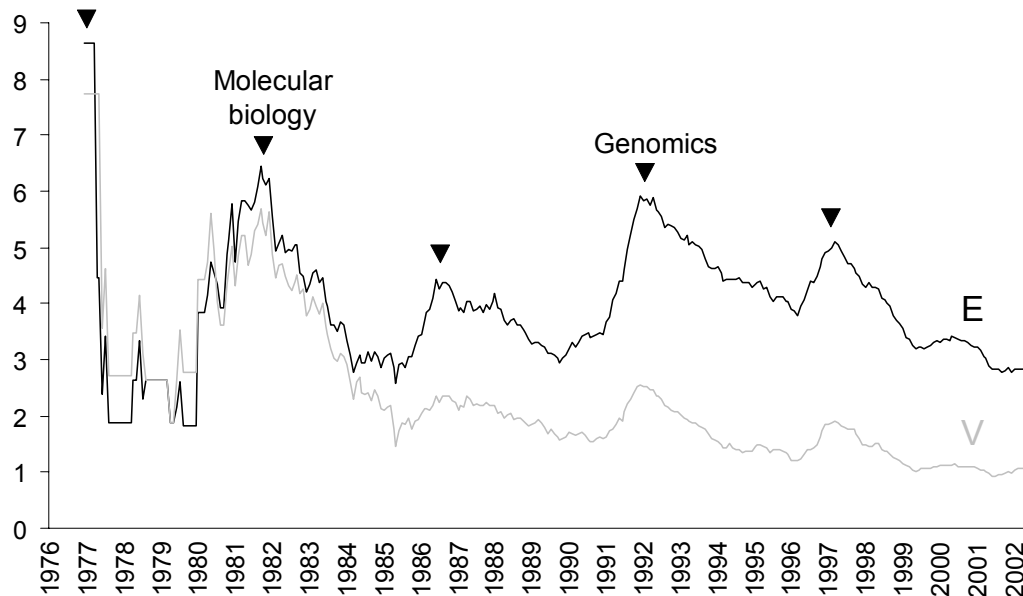
1.4 Outline of the book

The chapters of this thesis are organized as follows. In chapter two we introduce the reader to the developments in the European biopharmaceutical industry over the last decade. We empirically explore proximities between European biopharmaceutical organizations from two perspectives: the regional clustering of firms in space and the network of strategic collaborations between organizations. The aim of this chapter is twofold: first, we analyze the changing features of the main regional clusters of biopharmaceutical activity by type of organizations present, nature of activities, and degree of technological specialization. The second part considers the topological features of the European network of strategic alliances over time. Moreover, these topological features are compared to theoretical predictions about the optimal network topology for exploration and exploitation activities. Finally, chapter two concisely describes the preparation process of the data we used for chapter two and chapter four.

Chapter three focuses on the innovation strategies of the firm with respect to complementarities between in-house and external R&D. At the level of the individual research project, we identify the main determinants of successful research projects in drug discovery. More specifically, we examine how cognitive proximity between in-house research and external collaborations increases the probability of successful drug discovery. By doing so, this chapter analyzes how firms strategize on their portfolio of make, buy and ally transactions.

In chapter four we jump two aggregation levels higher from research projects within the firm to regional cluster of European biopharmaceutical firms. The aim in this chapter is to understand the determinants of local and non-local collaboration amongst geographically co-located firms. In other words, we test the antecedents of geographical proximity co-occurring with relational proximity. One of the main issues that this chapter addresses is the contradicting evidence that exists about whether exploration activities require geographical proximity or not. If geographical proximity matters we expect to see mainly clusters, which are networks of local collaboration. If geographical proximity is not required we expect clusters to act as nodes in a global (or European) network of collaboration.

Chapter five shows how radical technological change induces alliance formation of firms and how the structure of the network changes as a result of this. The Genomics revolution in the beginning of the nineties sets the stage for our analysis. In response to the radical innovations surrounding the human genome project, new firms enter the network of existing players. Being confronted with the potentially disruptive technologies of the newly entering firms, incumbent organizations face the challenge of obtaining access to the new technologies. As a result, a wave of strategic collaborations characterizes the industry. This allows us to test whether firms facing radical technological change form either local linkages or distant linkages in the network. Or, put differently, whether novel information is obtained through relational proximity or rather through relational distance. Finally, chapter 6 concludes. Table 1.2 provides an overview of the chapters of this thesis.

Table 1.2 - Outline of this thesis

Chapter		Proximity	Main research question	Unit of analysis	Methodology
1	Introduction				
2	Proximities in the European biopharmaceutical industry: regional clusters and networks	Geographical proximity & relational proximity	1 How do the main European clusters evolve of time? 2 How does the European network evolve over time?	1 Regional cluster 2 Firm	Explorative descriptive
3	Complementarities and the R&D boundaries of the firm	Cognitive Proximity	What are the main determinants of successful research projects in drug discovery?	Project	Binary logistic regression analysis
4	When clusters become networks: alliance formation in regional clusters	Geographical proximity & relational proximity & technological (cognitive) proximity	What determines Local & non-local link formation?	Regional cluster	Negative binomial regression analysis
5	Radical innovation and network evolution	Relational proximity	How does radical technological change affect alliance behavior and subsequent network structures?	Network	Explorative social network analysis
6	Conclusion & suggestions for further research				

Chapter 2   Proximities   in   the   European biopharmaceutical industry: regional clusters and networks[1]

2.1 Introduction

This chapter provides an empirical overview of biopharmaceutical innovation activities in Europe over the last decade (from 1996 until 2005). Our focus is on three forms of proximity between organizations in the European biopharmaceutical industry; geographical proximity, relational proximity and cognitive (technological) proximity. In our previous chapter we argued that these proximities help us understand firms' strategies in acquiring access to external knowledge, which is considered to be a prerequisite for innovation in the biopharmaceutical industry. In this chapter we empirically explore proximities between European biopharmaceutical organizations from two perspectives: the regional clustering of firms in space and the network of strategic collaborations between organizations.

We start with the identification of the main clusters of biopharmaceutical organizations in space, based on cluster (or regional) attributes on the one hand and on relational attributes on the other hand. Together these cluster-and relational attributes help us to identify certain 'types' of clusters, based on the literature of industrial districts (Markusen, 1996). We might for example find typical "hub-and-spokes clusters", where small firms evolve around academic centers of excellence (representing the 'hub'). Alternatively we might encounter clusters that resemble

---

[1] This chapter builds on a previous version (co-authored by G.A. van der Knaap), which is published in *Pharmareview*, June 2006.

"satellite clusters" where large pharmaceutical firms form mainly marketing and distribution alliances with non-local partners.

In the second part of this chapter we change our perspective towards a network view on innovation activities in the European biopharmaceutical industry. In order to do so, we 'map' all strategic alliances at three moments in time (1999, 2002, and 2005) as collaboration networks between organizations. We distinguish alliances as either science-driven or as market-driven. As a result the collaboration network is divided into an explorative- and an exploitative network. Lastly, we analyze some structural features of these networks over time and relate these to existing theory on innovation behavior of organizations.

Before we start exploring, we briefly explain how our data were collected and we define our variables of interest.

2.2 Data

2.2.1 Data preparation

Our information on European biopharmaceutical organizations, their relational activities, and their location originates from Pharmadeals alliance database between June 1996 and June 2005. Pharmadeals is a global monitor of alliances through press releases and annual reports search on a daily basis since 1996. As these concern alliance data, we are aware of organizations existing in Europe only if they have formed at least one alliance every three years. In the pharmaceutical industry, where external collaborations are deemed necessary for innovation, it is generally acknowledged that this selection criterion encompasses the majority of viable organizations in the industry (Arora & Gambardella, 1990).

The identification of the main clusters of European biopharmaceutical organizations has been a three stage process. To start with, we selected alliances where at least one European partner was involved. The database did not at that time allow us to create a query based on the location of organizations, and we used European countries as keywords to search for a match in any part of the alliance announcement (including press release). This resulted in a hit of 2800 alliances between June 1996 and June 2005.

In a second stage we have split alliances where multiple partners were involved. Of these 2800 alliances, some 300 appeared to involve more than 2 separate organizations. We decided to split those alliances into each possible combination of alliances. Thus, an agreement with *n* participating organizations was transformed into *n* x (*n*-1)/2 linkages. Hence, we assume that an alliance serves as a conduit of knowledge transfer, where information transfers between all participants in an alliance (whether the alliance is an R&D collaboration between multiple universities or a project of some pharmaceutical companies who received EU funding). Turning all these multiple partner alliances into dyads led to a new alliance set of 4031 alliances among 2500 separate organizations.

Thirdly we manually looked up the cities where organizations are located and found around 650 organizations either not tractable or undisclosed by Pharmadeals or being a USA based firm with a USA based partner (this latter phenomenon can be explained by the splitting of multiple partner projects). Our final work set consists of 2566 alliances among 1834 organizations of which 1054 are organizations located in Europe.

To identify the whole network of collaborations at various moments in time we created 'snapshots' of the network of collaborations in June 1999, June 2002 and in June 2005. These 'snapshots' build on the assumption of alliance duration of 3 years (Phelps, 2003), which means that each 'snapshot' captures all new agreements announced back to three years before the 'snapshot' was taken.

2.2.2 Cluster identification

Regional clusters of biopharmaceutical organizations have been identified by geo-coding each organization's location and using a hierarchal clustering algorithm based on Euclidian distances between locations to define the boundaries of European clusters. Having identified the European clusters we aggregated the information on organizations and their alliances back to the regional clusters. The following information per cluster was obtained:

Cluster name: city where most organizations are located

Cluster size: number of organizations that have formed at least one alliance in the past three years.

Alliances: number of alliances formed by organizations located in a cluster. These alliances encompass all phases in the drug development process ranging from discovery collaborations through licensing deals to marketing-, and distribution alliances. Furthermore, we identified all alliances as either local collaboration with both organizations located in the same cluster, or as non-local collaborations whereby collaborating actors are located in different clusters.

Company type per cluster. Our data contained information about the type of organizations that are involved in each alliance. This enabled us to identify the number of start-up firms, large established organizations, governmental organizations, academic organizations and finally some financial firms per cluster.

Therapeutic focus of the alliances of organizations located in each cluster. Each alliance focuses on a specific therapeutic area or on a combination of therapeutic areas. In some cases the alliance is not therapeutically focused but rather focuses on a technology such as genomics, proteomics or on certain platform technologies such as high throughput screening or assay detection. In any case, in line with the relational view of the firm, we consider the whole portfolio of activities in which a firm is deal-active as the resources of the firm (Dyer & Singh, 1998). As such, the collection of activities of all firms in a cluster can be considered as the resources of a cluster.

Finally, our database contains information on the phase in the drug development process at which each alliance is targeted. These phases range from discovery and lead optimization to clinical testing and lastly marketing and distribution activities. We have exploited this information to disentangle explorative activities from exploitative activities and subsequent network structures.

2.3 Geographical proximity: regional cluster development

Most of the 1054 biopharmaceutical organizations in Europe tend to co-locate in a few regional clusters. In fact, our data show that 72 percent of the active biopharmaceutical organizations are located in the 30 largest regional clusters. While the remainder of this chapter focuses on the largest clusters, we first summarize cluster characteristics over time for *all* regions (where at least 2 deal-active organizations are clustered) in table 2.1.

Table 2.1 – European biopharmaceutical clusters in 1999, 2002 and 2005 (descriptive)

	Jun-99	Jun-02	Jun-05
Number of organizations	261	380	568
Linkages	942	956	1679
Average number of linkages per firm	3.6	2.5	3.0
Mean number of organizations per cluster	5.44	6.23	7.78
Skewness of distribution of organizations	1.69	2.96	2.84
Mean number of linkages per cluster	19.63	15.67	23.00
Skewness of distribution of linkages	2.59	3.05	3.44

In the last decade there has been a steady stream of new clusters of biopharmaceutical organizations in Europe. As a result, the total number of active organizations in these clusters has increased. Although the total number of alliances has risen, the average number of alliances per firm reveals a more cyclical trend. The average organization has become less deal-active in the first period between 1999 and 2002, while it has become increasingly deal-active in the second period between 2002 and 2005. The cyclical nature of link formation is a common feature of alliance behavior among firms (see Hagedoorn (2001)). When we look at the mean number of organizations located in a cluster, we see that on average clusters have been growing over time. This average growth is remarkable considering the fact that in each period there are 10 new (and usually small) clusters emerging on the European scene. It indicates that the growth in the number of active organizations has been larger than the growth in the number of new clusters. Lastly, we consider the distribution of clusters in terms of number of organizations over time. In table 2.1 we see an increased skewness in the distributions of cluster size in the period 1999 – 2002, which indicates an increased inequality in the number of organizations per cluster. In other words, in the first period the number of larger clusters has been growing faster than the number of smaller clusters. In the second period, from 2002 till 2005, we see a decreased skewness in the distribution of cluster size, which indicates an increased equality in

the number of organizations per cluster. Whereas agglomeration effects seem to be at work in the first period (resulting in a large-grow-larger phenomenon), the second period is characterized by a (light) catching up of smaller clusters in terms of number of organizations. In sum, table 2.1 reveals that for organizations in the European biopharmaceutical industry both link formation and co-location (or clustering) are not activities that grow linearly over time, but are cyclical. Technological changes are often said to underlie these cyclical patterns in link formation and clustering behavior of firms (Tushman and Anderson, 1986; Nooteboom, 2001). If both link formation and regional clustering are part of a firms' strategy to gain access to external knowledge, then one could argue that in times of more radical technological change, the necessity for firms to access external knowledge increases and hence link formation (related to R&D) and regional clustering increases. In chapter 5 of this thesis we empirically investigate these hypotheses in detail. More specifically, we study the effect of radical technological change on patterns of link formation in the pharmaceutical R&D network.

2.3.1 Identifying the main clusters and the main activities.

Which are the main clusters of European biopharmaceutical firms? Where do biopharmaceutical organizations locate and which activities do they employ? To answer these questions we focus on the 30 largest clusters in terms of the number of organizations present. Within these clusters we distinguish between different types of organizations and different therapeutical areas in which organizations are active.

Main type of organizations per cluster

Figure 2.1 shows the 30 largest clusters in the European pharmaceutical industry by type of (deal-active) organization over time. We distinguish among academic organizations, biotech firms, and (large) pharmaceutical firms. Governmental organizations and financial firms are important actors in the innovation process, but they are hardly captured by our data of (commercial) strategic alliances.

Fig. 2.1 — 30 main clusters by type of organization in 1999, 2002, 2005
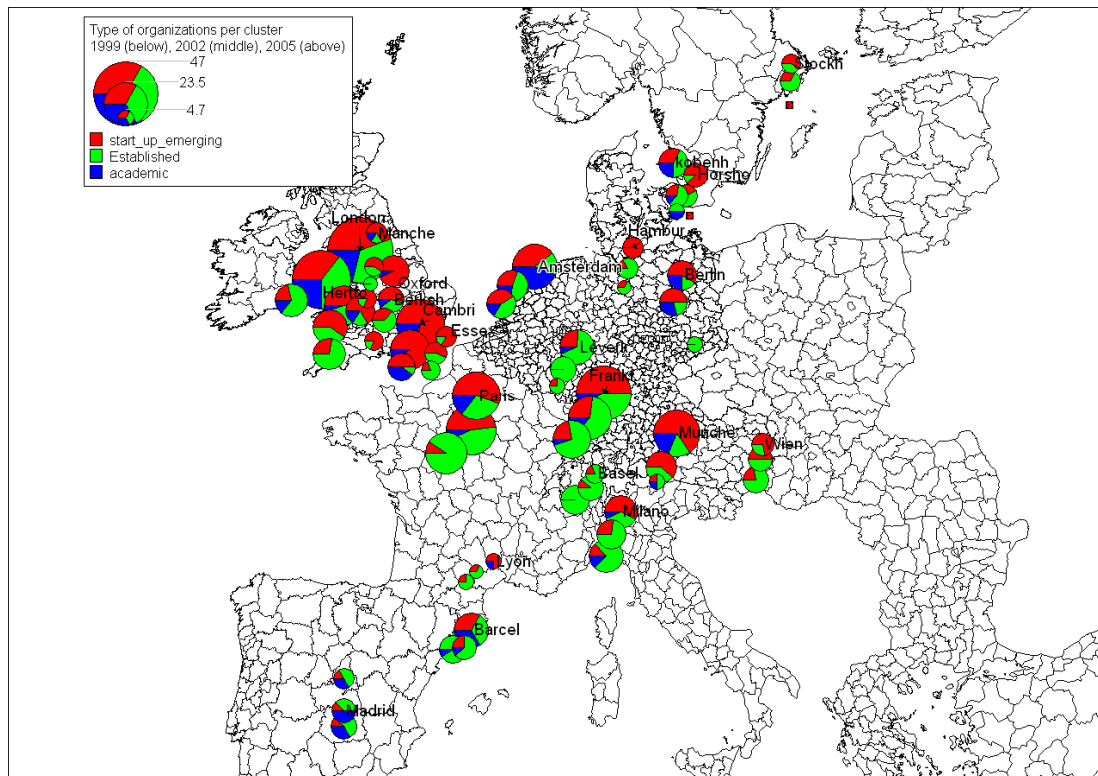
European pharmaceutical clusters are increasing in size between 1999 and 2005. In the same period we see a growing dominance of London and surrounding clusters (including Essex, Cambridge, Oxford and Berkshire) as the largest pharmaceutical cluster in terms of deal-active organizations present. One might even typify the area around London as one giant cluster of biopharmaceutical activity, since 20 percent of all European biopharmaceutical organizations (211 organizations) are located in this area. This visual observation is confirmed as a general trend of increasing inequality in the distribution of organizations over European clusters (standard deviation of organizations per cluster more than doubles from 1999 until 2005). Indeed, clustering or geographical proximity is an increasingly important phenomenon in the European biopharmaceutical industry. The clusters that surround London can be divided into clusters that are driven by large pharmaceutical firms (Hertfordshire, Manchester, Essex) or driven by academic organizations such as Cambridge and Oxford. In both cases the number of start-up firms has gradually increased over time. For the rest of Europe, only the Madrid cluster has had a steady majority of academic organizations involved in alliances in the last decade. The proportion of start-up firms has increased

in almost every European cluster, with the extreme cases being Horsholm, Munich and Vienna. Finally Paris and Frankfurt stand out as the clusters with a majority of large established pharmaceutical organizations present.

Main activities per cluster

Therapeutic focus

Activities in drug development are often categorized along their therapeutic focus[2]. Usually a therapeutic focus contains various disease areas and can be approached using different technologies. For example, Immunological is a therapeutic focus area which covers a number of diseases related to the immune system, such as rheumatoid arthritis or multiple sclerosis. Compounds to cure these diseases can be either chemically based or biologically based and the technologies that are available to screen, select and test these compounds are numerous. Since the beginning of the nineties, the genomics revolution has brought a set of technologies that aid the drug development process by being applicable to any compound without needing a specific therapeutic focus. As a result, alliance activities can be either focused on a specific therapeutic area or on a genomics related technology such as genomics, bioinformatics, high-throughput screening etc. For visualization purposes, we have grouped alliance activities around basic research, diagnostics, drug delivery, and around genomics related technologies for each cluster over time. Figure 2.2 visualizes the main activities per cluster.

---

[2] With the exception of anti-cancer therapies. This is such a broad disease area that it has become a therapeutic focus area on its own.

Fig. 2.2 — 10 main clusters by type of activity in 1999, 2002, 2005
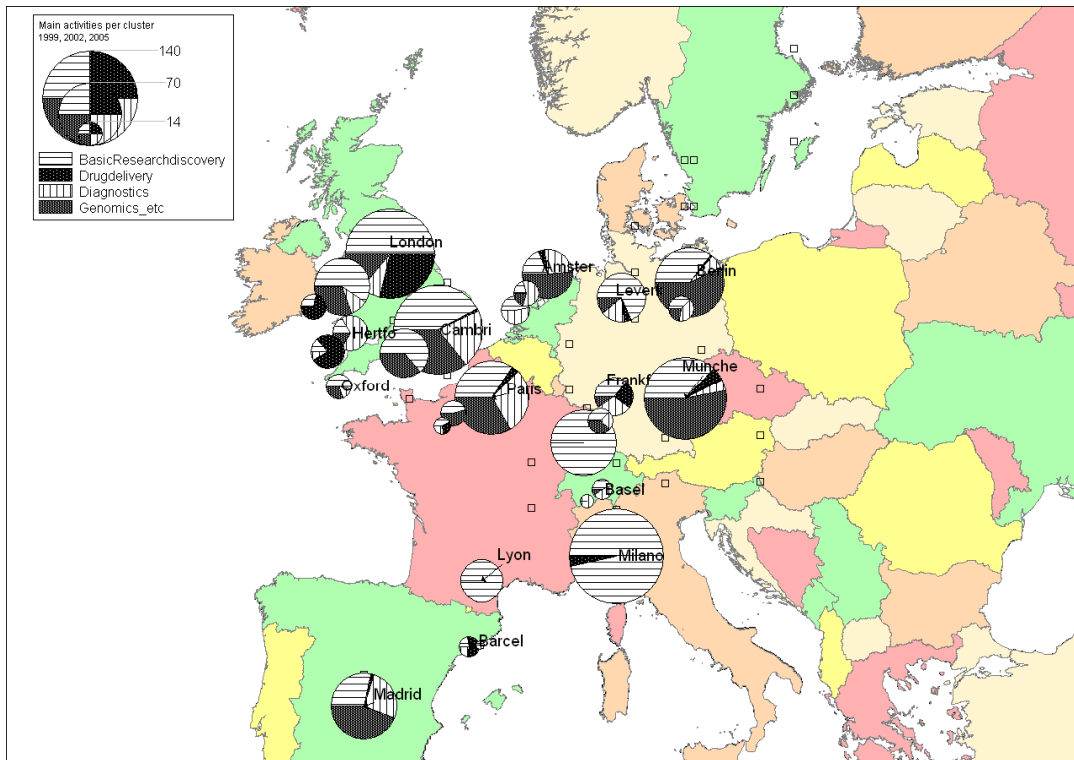
Basic research, diagnostics and drug delivery can be seen as activities that take place at the beginning, the middle and the end of the drug development process. Genomics related technologies are so-called 'general purpose' technologies that are used as instruments to enhance the drug development process without being part of it. In figure 2.2 we can see that basic research is an important part of all European clusters. More interestingly, it seems that clusters which are dominated by either start-up firms (such as Cambridge Munich and Berlin) or by academic organizations (such as Madrid and Amsterdam in 2005), a higher proportion of activities is geared at genomics related technologies. This finding can be explained by the fact that the genomics revolution in the beginning of the nineties was sparked by global academic research and caused a number of waves of new firm entrance around the introduction of new genomics based technologies such as genomics, proteomics or bioinformatics.

Regional specialization

Based on the therapeutical or technological focus of alliances we created a measure of regional specialization. More specifically, we consider the portfolio of alliances of

organizations in a cluster as the resource of the cluster. Using the Herfindahl index enables us to determine to what extend a cluster is dominated by a relatively few therapeutical (or technological) areas. Originally the Herfindahl index is a measure of the size of firms in relation to the industry and an indicator of the amount of competition among them. In our case the Herfindahl index expresses the degree to which a biopharmaceutical cluster is dominated by a small number of therapeutic areas/technologies. The Herfindahl index is defined as follows:

$$\sum_{i=1}^{k} \left[\frac{N_i}{\sum_{i=1}^{k} N_i} \right]^2$$

The index is the sum of the squared share of each therapeutic area in a cluster. N_i is the number of deals that are based on therapeutic area i. There are k therapeutic areas and N deals in a cluster. The reason for squaring each share of therapeutic areas is to put more weight on larger therapeutic areas and less weight on smaller areas. We use the Herfindahl index to typify clusters as being more or less specialized in certain therapies or technologies in table 2.3. The exact therapeutical areas in which clusters are specialized are visualized in figure 2.3

Fig. 2.3 – Therapeutical specialization of the main
European clusters
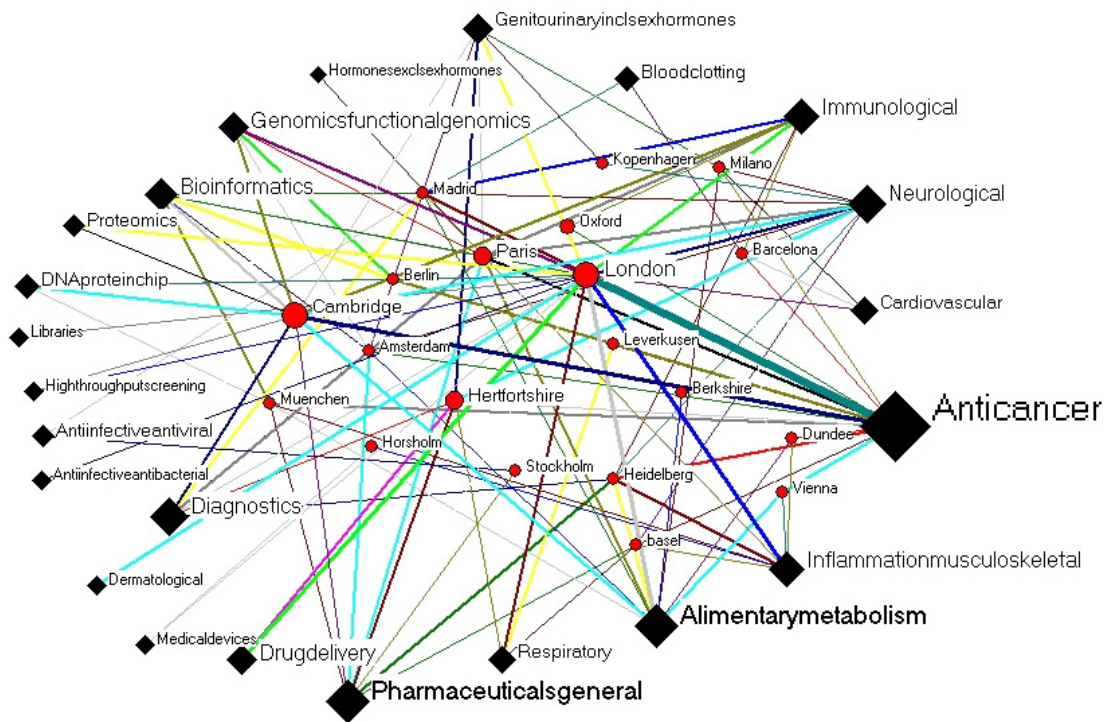
In figure 2.3 we have accumulated the therapeutic areas in which organizations in each cluster close many deals (10 deals or more). This leads to a network of clusters and therapeutic areas whereby the thickness of the ties represent the number of deals closed by organizations in a cluster based on that specific therapeutic area. The size of the therapeutic area nodes represents the number of times that the therapeutic area is dealt by any European organization, and the size of the cluster nodes is determined by the total number of deals closed by organizations in the cluster.

A number of observations can be made from this network at first glance. First, anticancer is by far the most popular therapeutic area in Europe. Almost every big cluster has closed at least ten deals based on anticancer therapies. The latter observation however is also valid for other therapies such as neurological, immunological and general pharmaceuticals. The difference between anticancer and the other mentioned therapies lies in the extraordinary strong relation between London and anticancer therapies. In fact, 142 deals have been closed in London based on anticancer therapies. While London appears to be strongly focused on its main disease area we will come to see later, when we calculate the degree of therapeutical

specialization for each cluster, that this strong focus does not make London a particularly specialized cluster. Because of its large size London is an example of a mixed case with a number of specialized sub-clusters and at the same time a wide range of different types of therapeutics. Our final observation is about the similarities in alliance activity between Cambridge and London. From the above figure we see that they have closed about the same amount of deals from 1996 till 2005 (node size). Cambridge ranks second in terms of number of organizations after London (figure 2.1), and while it is deal-active in fewer different therapeutic areas, it has a similar (though weaker) dominant tie with anticancer therapeutics as London has.

2.3.2 Positioning of main clusters in the European network: relational attributes

Where do organizations in a cluster search for external knowledge? Do they search locally within their cluster or do they search for external knowledge *beyond* their cluster? By answering this question we do not only gain information about the relational behavior of organizations in clusters, we also obtain insights on the European network of inter (and intra) regional relations in the pharmaceutical industry and the position of clusters in this network.

Exploration network & exploitation network

When studying the European network in terms of local and non-local search activities, we distinguish between explorative search and exploitative search. Our data capture two types of agreements between organizations: exploration alliances that are aimed at exchanging knowledge or technologies in the earlier stages of drug development, and exploitation alliances aimed at downstream drug development (see data section for a more detailed description). Exploration alliances are known to involve relatively

more tacit knowledge in comparison to exploitation alliances[3]. It has been argued in the literature that tacit knowledge transfer is more likely to require geographical proximity than explicit knowledge transfer (Pavit, 1987; Audretch & Feldman, 1996). As a result we expect to see explorative alliances to occur locally more often than exploitation alliances.

Fig. 2.4 – Network of exploration activities in the European biopharmaceutical industry in 1999, 2002, and 2005.

Node size represents the relative importance of local collaboration and ties represent the relative importance of non-local collaboration
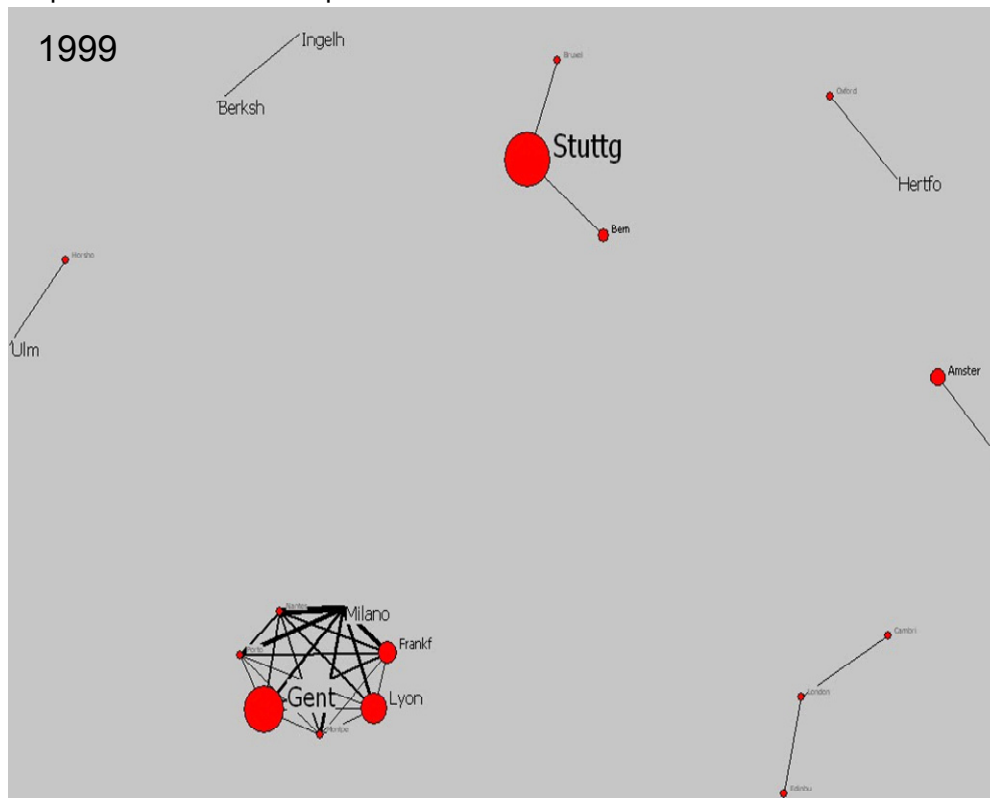
---
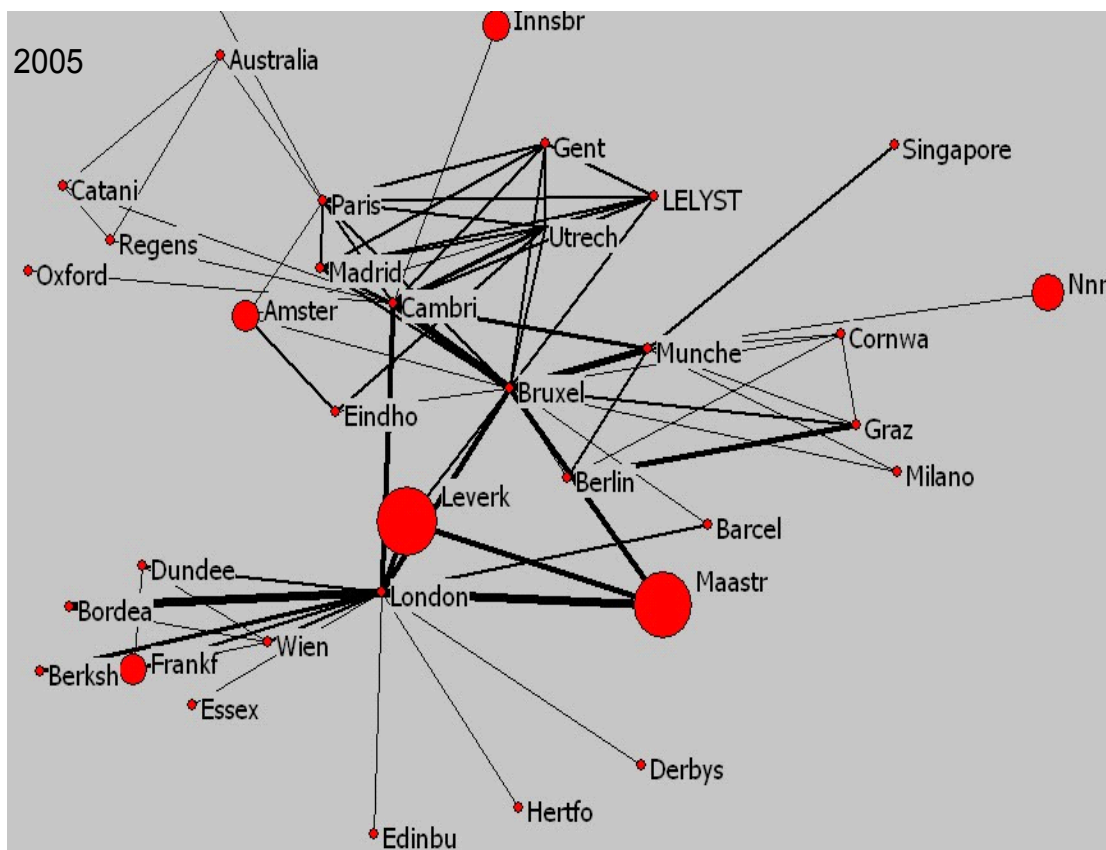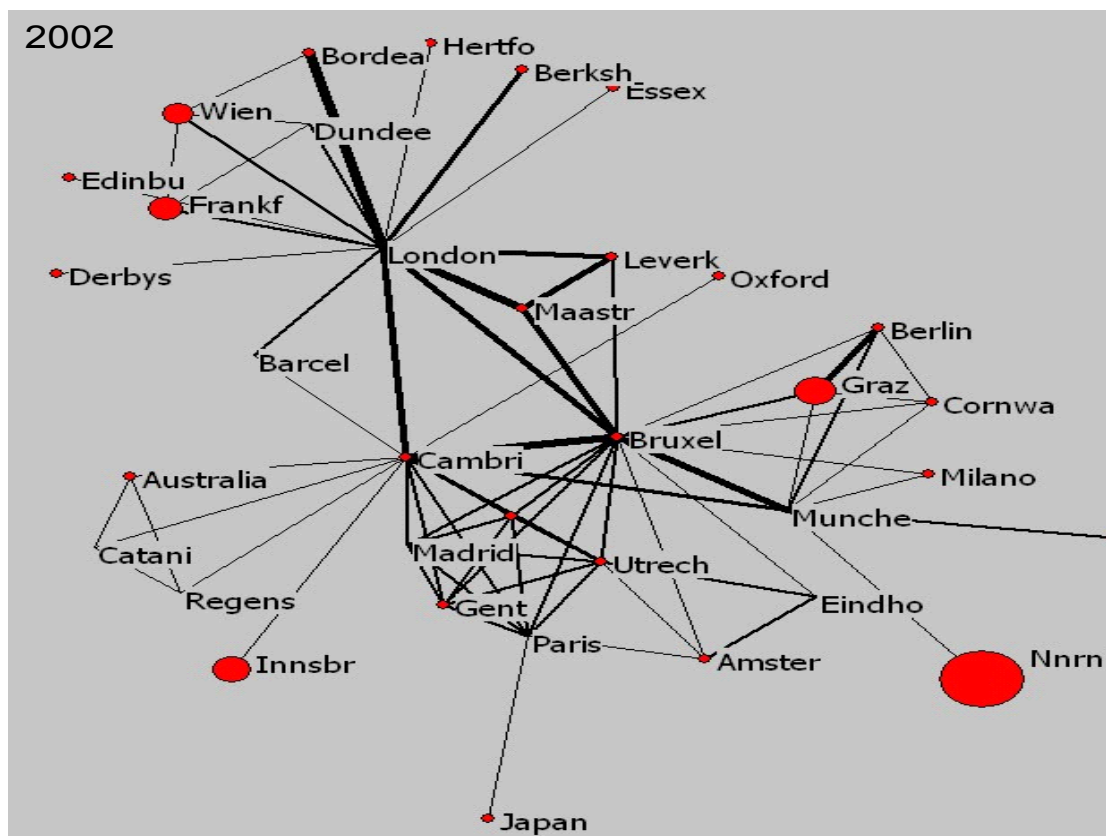
Fig. 2.5 – Network of exploitation activities in the European biopharmaceutical industry in 1999, 2002, and 2005.

Node size represents the relative importance of local collaboration. Ties represent the relative importance of non-local collaboration
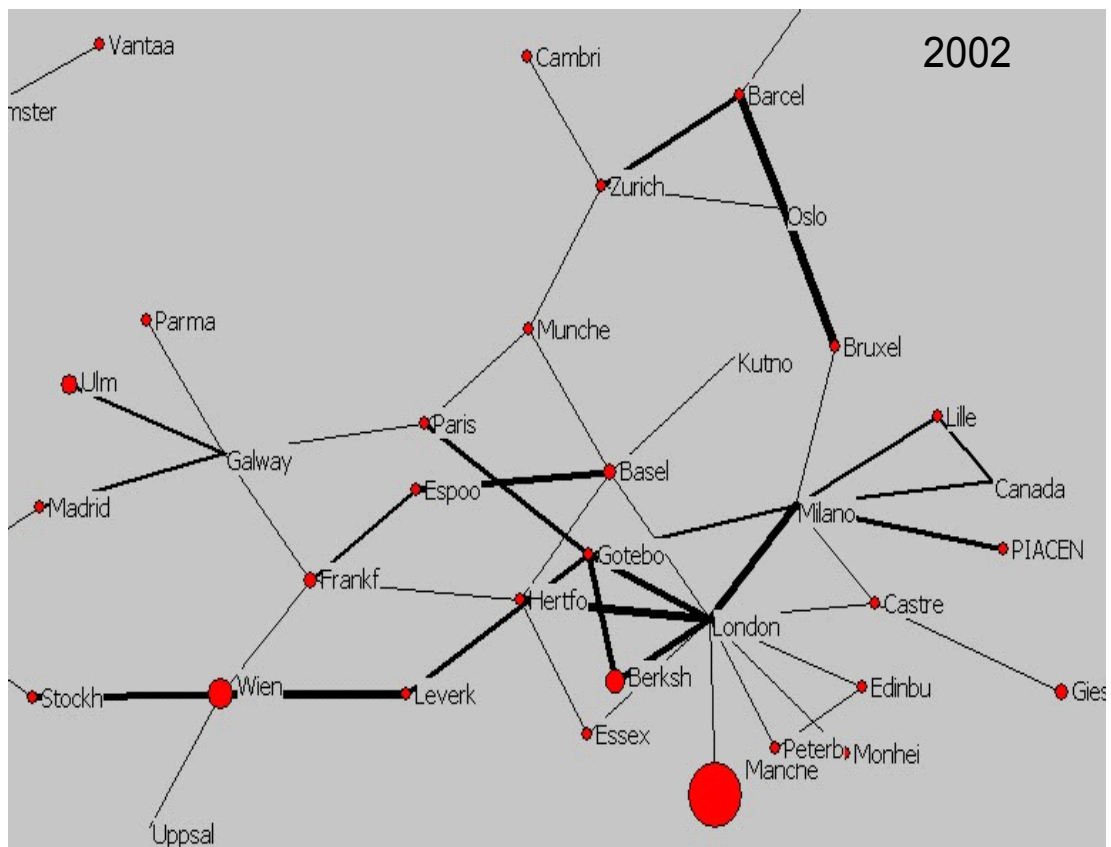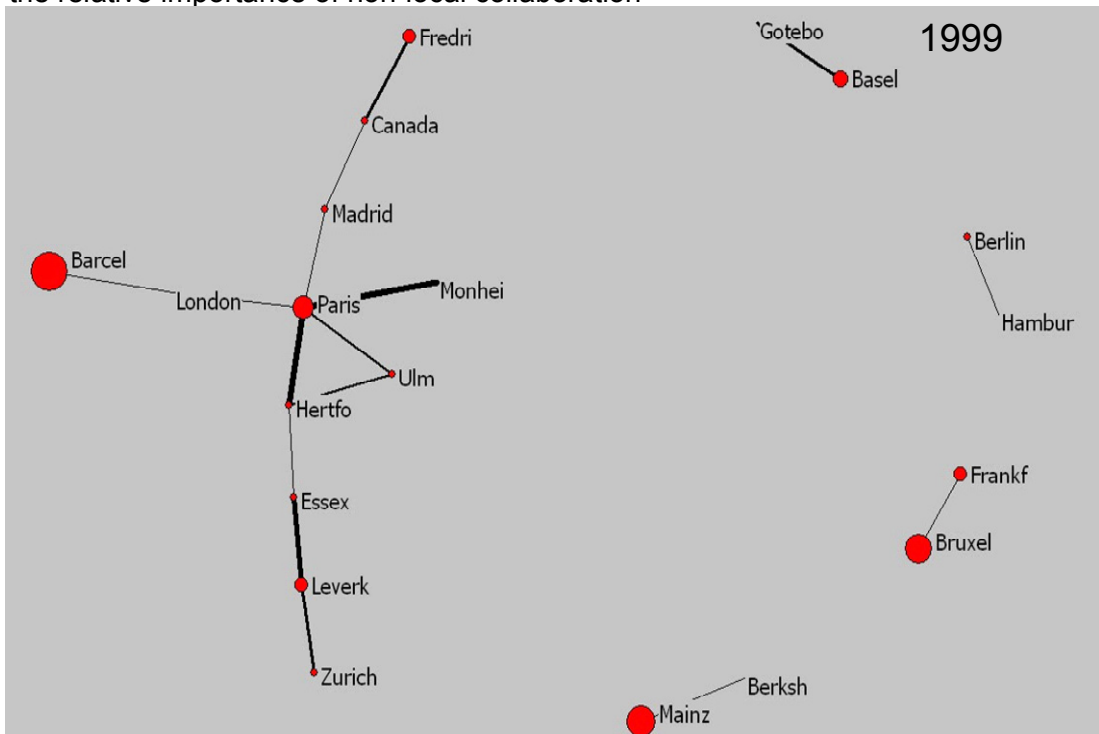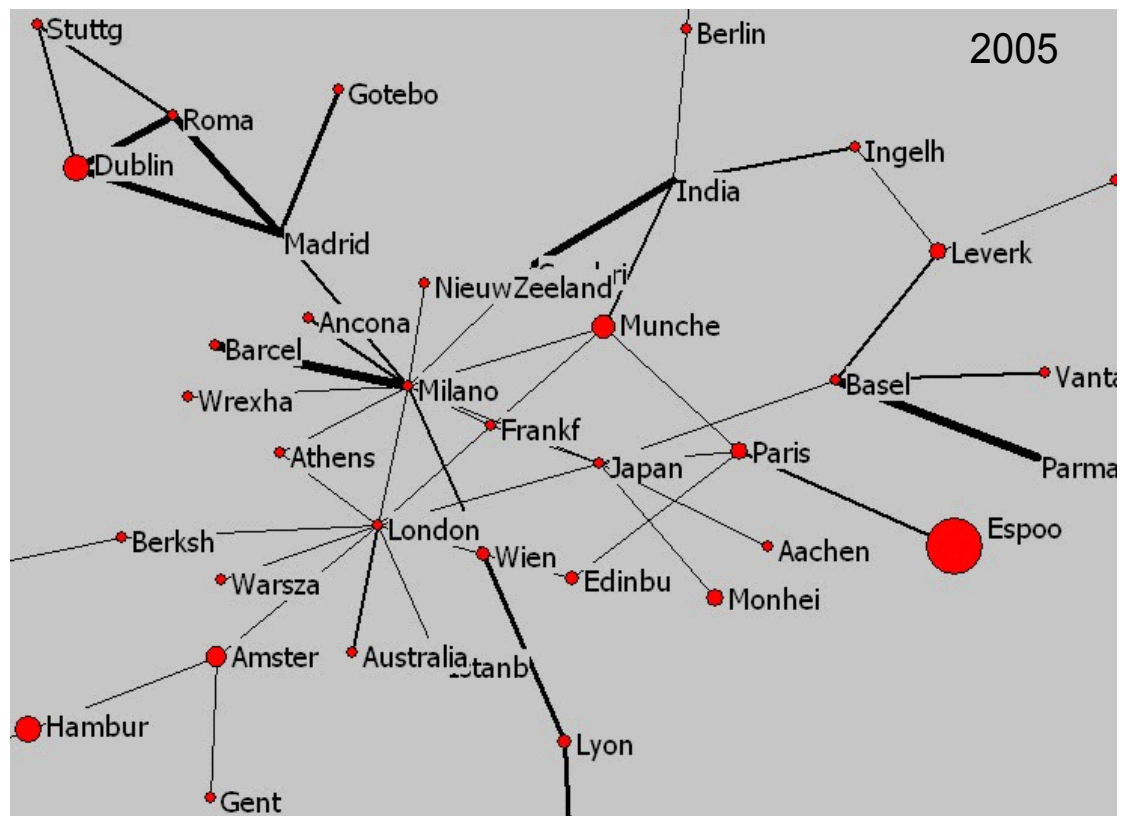
Table 2.2. - Local & non-local collaboration in exploration & exploitation networks

	Exploration			Exploitation		
	1999	2002	2005	1999	2002	2005
Local links	29	18	49	2	2	17
Non-local links	287	164	603	188	228	255
Local/all links	0.09	0.10	0.08	0.01	0.01	0.06

Figures 2.4 and 2.5 represent the exploration and the exploitation networks over time. Node size represents the *relative* importance of local linkages while the relations between nodes represent inter-cluster linkages. Table 2.2 summarizes the information in the figures 2.4 and 2.5 over the main 30 clusters and provides information on the *absolute* number of local and non-local alliances over time. From the table we can see that at any point in time non-local linkages occur at least ten times more often than local linkages. This finding might be caused by the fact that our data capture only alliances above a certain deal value. If it is the case that organizations form local alliances as 'informal' or 'supportive' alliances rather than as formal, commercial deals, than this could explain the very low amounts of local alliances in our data.

Furthermore, when we consider the amount of local linkages in both networks in table 2.2, we see that explorative activities are much more often local than exploitative activities[4]. This finding is in line with our expectations from the literature stating that explorative activities contain more tacit knowledge which requires geographical proximity.

In figure 2.4 we see how the exploration activities, such as R&D collaborations and drug discovery efforts are being undertaken by European organizations at three points in time. In 1999 there is a remarkable clique of mainly southern European clusters that are strongly dominating the European scene. The clusters in this clique (being Milan, Frankfurt, Lyon, Nancy, Porto and Gent) are fully connected[5], which is caused by the fact that the organizations located in this clique were participating in one research project subsidized by the European Union (the so-called Eureka Peptido project). In 2002 and 2005 it becomes clear that the main explorative activities have shifted to London, Brussels and Cambridge (with minor roles for Maastricht, Leverkusen and Munich and Madrid)

---

[4] Even the share of local alliances in relation to all alliances is higher in the exploration network than it is in the exploitation network at any point in time.

[5] Indicating that in a clique with N actors, there are $N *(N-1) / 2$ linkages between actors.

In figure 2.5 we see how the exploitation activities, such as marketing and distribution activities are being undertaken by European organizations at three points in time. In 1999, Paris was performing a central role in the exploitation of existing drugs in Europe. From 2002, London has again (like in the exploration network) taken over the lead position in the European exploitation network with Milan close behind. Interestingly, in 2005 Milan has gained the dominant position in the network, leaving London at a second place. The activities performed by organizations located in Milan have shifted strongly from the early phases of drug development (R&D and discovery) in 1999 toward late stage drug development (marketing and distribution) in 2002 and 2005. While beyond the scope of this study, it might be worth-while to consider whether these changes in Milan represent cluster life cycle effects.

2.3.3 Types of clusters

Thus far, clusters have been identified in terms of the type of organizations present, the therapeutic focus of activities and in terms of local and non-local search activities. In order to see whether there exists certain 'types' of clusters based on the above and other attributes, we have sorted the main clusters along their degree of therapeutical specialization and added a number of other attributes[6]. In the following table we see the 3 most specialized clusters (Lyon, Milan, Horsholm) and the 3 most diversified clusters (London, Paris, Amsterdam) with regard to therapeutic focus. Besides the identification of these clusters by name we use the cluster characteristics as shown in the columns to describe the clusters.

---

[6] Alternatively, we tried to identify cluster 'types' by using factor analysis of cluster attributes over our main (30) clusters. This did not lead to any cluster type.

Table 2.3. — Most specialized and least specialized European biopharmaceutical clusters and their attributes

cluster name	Herfindahl	Organiza-tions	Nodes	Start-up ratio	Openness cluster*	USA connectivity*	R&D ratio
Lyon	0.26	28	10	0.43	1.06	0.41	0.74
Milano	0.24	65	31	0.38	0.92	0.66	0.52
Horsho	0.16	38	15	0.67	0.66	1.16	0.09
London	0.07	297	86	0.58	0.92	1.13	0.33
Paris	0.06	153	67	0.44	0.98	1.32	0.2
Amster	0.06	58	52	0.59	1	1.64	0.34

*These two cluster characteristics are expressed in relation to the average score in all European clusters. The scores range from 0 to 2, whereby 1 indicates the average score of all European clusters.

The Herfindahl index shows that the clusters Lyon, Milan and Horsholm are on average four times more specialized than London, Paris and Amsterdam. Further it seems that the bigger clusters (number of organizations) are the more diversified ones, but simple correlation statistics of all clusters reveal this is not the case. What does matter for the specialization of a cluster is the share of deals that are closed with a partner from the USA. More diversified clusters have relatively more partners in the USA than specialized clusters have.

As not all organizations located in a cluster are deal-active, we have added a separate column for deal-active firms. It is interesting to see that while for all clusters but Amsterdam (Amsterdam, Leiden, Rotterdam region) around 35% of the organizations are deal-active, 90% of the Amsterdam organizations are deal-active. Taking a closer look at the Amsterdam cluster we see that it is not only the most active cluster, it is also very 'USA oriented'. The Amsterdam cluster has over one-and-a-half times more connections to the USA than other European clusters have.

A second cluster that deserves extra attention is Horsholm in Sweden. Horsholm seems to be a typical neo-Marshallian district in that it is one of the most specialized clusters in Europe, in combination with a relatively large share of start-up firms who

are closing deals mainly to partners inside the Horsholm cluster. Further it seems that Horsholm is specialized in downstream drug development because of its few R&D collaborations.

Lastly, there is the Lyon cluster which is rather exceptional. Lyon is the most therapeutically specialized cluster among Europe's main clusters, and on top of that it has a very strong focus on R&D together with a relatively weak connection to the USA. Lyon can therefore be typified as a specialized R&D cluster.

2.4 Relational proximity: inter-firm network development

To assess empirically the importance of relational proximity as a governance instrument for accessing external knowledge, we take a look at the networks of collaborative agreements among European biopharmaceutical organizations over time.

As discussed previously, organizational search for external knowledge is usually divided into either local or non-local search (Rosenkopf & Nerkar, 2001). In paragraph 3.2 of this chapter we have seen from a geographical perspective that firms in a cluster rather search beyond the boundaries of a cluster than within the cluster. Regarding relational proximity, both local and non-local search can be regarded as a prerequisite for knowledge access and innovation. To understand how, we shortly explain local and non-local search in network terms, and how they are associated with innovation.

2.4.1 Local & non-local search in the network

Local search manifests itself through local linkages. In network terms, a linkage is local if it is embedded in a relatively dense web of previous and neighboring relations. Together this dense web of previous and neighboring relations form a local clique of collaborating and knowledge exchanging organizations, which through imitation behavior subsequently converge toward cognitive proximity (Boschma, 2004). Non-local search in a network leads to a distant linkage (sometimes referred to as a 'weak' linkage), which means that this link forms a bridge between two actors (or cliques) in

a network that would ordinarily have to surpass a large number of actors to reach each other (Newman, 2001). Just as local linkages are often assumed to carry more similar knowledge (cognitive proximity) it is also assumed that distant linkages correspond to high cognitive distance between actors, whom bring together new pieces of knowledge (Rosenkopf & Nerkar, 2001).

2.4.2 Local & non-local search & innovation

In line with the transaction costs view of the firm, which emphasizes mainly the costs and threats of external knowledge acquisition, Cowan (2005) argues that for the exploration of new knowledge and new combinations, local search is more beneficial than non-local search. In exploration, Cowan argues, knowledge is rather tacit and therefore difficult to absorb. The difficulty of absorption capacity can be overcome through the formation of local linkages in dense cliques where knowledge domains are more similar. Together the clique members are better able to create new knowledge and to innovate. Once a dominant design has emerged and is ready to be exploited, knowledge becomes more codified and hence easier to absorb. At this stage the organization aims to diffuse its innovation as much as possible and non-local search through distant linkages are most optimal to do so.

When authors apply the resource based view of the firm, as Sidhu, Commandeur and Volberda (2007) do, the innovation strategy that is considered optimal reverses. These authors argue that in order to explore in a dynamic environment, firms are required to reach out to organizations that are relatively far away from their current field of expertise, so as to make novel combinations. Local search through local linkages, which provide access to similar knowledge, enable incremental improvements necessary for the exploitation of existing products. The main difference between these two studies lies in the perceived importance of a firm's absorptive capacity. Whereas Sidhu et al (2007) emphasize that the gains from novel combinations (through non-local linkages) outweigh the costs of integrating them with the internal knowledge base, Cowan argues that the costs of absorbing knowledge requires local embeddedness. In addition to these arguments, a number of other issues might play an important role. Regarding local search and local linkages, there might be costs associated with redundancy or from unintended knowledge flows, especially if relational proximity co-occurs with cognitive proximity (knowledge similarity). With

respect to non-local linkages, not only the high integration costs of new knowledge might be problematic, but also the danger of opportunistic behavior could be higher since a distant linkage cannot rely on the protection of reputation effects in the network. Finally, the latter potential danger might also turn into a negotiation advantage if this distant linkage appears to be bridging two cliques with complementary knowledge. The intriguing question becomes what we observe empirically in an extremely dynamic environment as the European biopharmaceutical industry, where explorative relational activities can be clearly separated from exploitative relational activities.

2.4.3 Local linkages and non-local linkages in exploration and in exploitation networks

Whereas in paragraph three on geographical proximity we explored our data at the level of the geographical cluster, in this section we analyze relational proximity (through local and non-local linkages) at the level of the network. Again we compare these networks over time and we look at explorative and exploitative networks separately. We start with some descriptive characteristics of our networks and turn our attention to measuring local and distant linkages.

Table 2.4 – Comparing structural network features of the exploration and exploitation network over time

	EXPLORATION			EXPLOITATION		
	1999	2002	2005	1999	2002	2005
Clustering coefficient (c)	1.933	0.232	0.268	0.000	0.000	0.000
Average path-length (d)	1.862	2.898	5.416	5.304	2.823	3.246
Number of organizations (giant component size)	120 (18)	211 (20)	433 (193)	130 (35)	177 (19)	235 (26)
Number of linkages	287	164	603	188	228	255
Average degree	2.39	0.78	1.39	1.45	1.29	1.09

While both the exploration and the exploitation networks are growing over time (in number of active organizations), it is becoming sparser as the average number of alliances per firm decreases in both networks. Particularly in the exploration network there is a sharp decrease between 1999 and 2002 in the number of alliances formed while the number of organizations has grown. If we compare these results with figure 1.3, which plots R&D alliance activity on a global scale, we can see that European organizations are behaving in line with global trends. Indeed, the wave of global alliance growth in figure 1.3 between 1996 and 1999 (induced by the entry of proteomics technologies) is fully captured by the 1999 picture in table 2.4 (as the 1999 network is a snapshot of newly announced alliances between 1996 and 1999). The decreasing number of global alliances formed in the 1999-2002 period is also in line with our findings of the 2002 network in table 2.4.

Based on the literature discussion above, our question at hand is the following: are local linkages (relational proximity) predominant in exploration networks while more non-local linkages characterize the exploitation network or vice versa? Underlying this question is the issue of whether firms develop knowledge search strategies based

on avoiding transaction costs or based on opportunities to create novel combinations. Local linkages together form local cliques in which everybody is connected to everyone else. The importance of the alliance behavior is typically measured by the clustering coefficient. The clustering coefficient of an actor is the density of its open neighborhood that is to say how close each actor's neighborhood is to a fully connected clique. Following Watts and Strogatz (1998) we define a clustering coefficient as follows: assume that the ith vertex v_i has $k_i - 1$ neighboring vertices. At most, $k_i(k_i - 1)/2$ edges can exist between them. Calculate $c_i \equiv$ (number of edges of v_i and its neighbors) $/ k_i(k_i - 1)/2$. We define the overall clustering coefficient as:

$$CC = \frac{1}{N} \sum_{i=1}^{N} c_i$$

CC is the average of the individual clustering coefficients c_i. The weighted overall clustering coefficient (*WCC*) is the weighted mean of the clustering coefficient of all the actors each one weighted by its degree.

Non-local linkages are measured by the average path-length or sometimes referred to as the diameter of the network. Non-local linkages function as bridges between different parts of a network and thereby they reduce the average distance between actors. We formally define average path-length as follows:

$$d(N) = \frac{\sum_{j \in V} \sum_{i \in V} d(i, j; N)}{v(v - 1)}$$

The average path-length *d(N)* is the average distance between any actor *i* and *j* that belong to the same component (giant component). Actors that are isolated from other actors are excluded. Thus, for a connected graph *N(E,V)* consisting of edges (linkages) and vertices (nodes), the average path-length is the sum of the distances between two actors (*i* and *j*) belonging to the network *(N),* divided by all possible edges excluding self-ties.

Table 2.4 reveals the clustering coefficient and the average path length over time for both networks. It becomes clear from this table that local linkages are a much more important phenomenon in exploration networks than in exploitation networks. This indeed indicates as Cowan (2005) has argued that new knowledge creation mainly takes place in local cliques when knowledge is still rather tacit and hard to absorb. The absence of clique formation in the exploitation network might indicate that more codified knowledge (which is transferred in exploitation networks) is preferably not exchanged in cliques to prevent unintended knowledge flows to other actors (e.g. competitors) within the clique. The extraordinary high clustering coefficient (1.93) in 1999 can be explained by taking a look at the network visualization of 1999 in figure 2.4 and figure 2.6. Although the nodes in figure 2.4 represent clusters and not firms, it is firms which form these alliances (see figure 2.6). Indeed the clique dominated by organizations in Milano and Gent is a full clique in which all clusters are connected.

With respect to non-local linkages, we can say that the more predominant they are, the shorter the average path-length will be. An average path-length can only be calculated for actors that belong to the main component of a network where every actor is reachable. Except for the exploration network in 2005, both the exploration network and the exploitation network consist of very few actors belonging to the main component, which makes the interpretation of the average path-length dubious. While there is a strong growth of deal-active organizations in both networks, for the most part these organizations remain active in independent components without forming an integrated European network. Interestingly, in 2005 European organizations which are active in exploration have managed to connect these previously unconnected components to a large extend. As a result, the 2005 exploration network shows a relatively large giant component. To put it differently, in 2005 (almost) every second European organization is connected to one giant European network of firms active in exploring new drugs. Figure 2.6 shows how the separate components in the exploration network become connected in 2005.

Fig. 2.6 – The emergence of a giant component in the (firm level) exploration network

1999

2002   2005
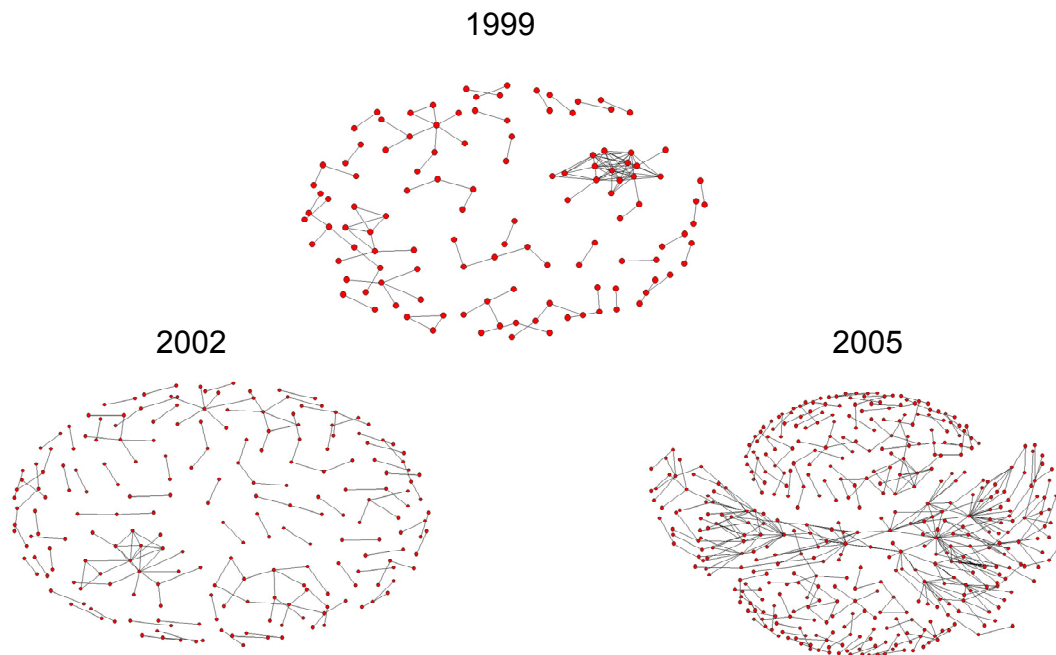
While we cannot rely on the average path-length information to assess whether non-local link formation has occurred, we can analyze local link formation by visualizing the emergence of a giant component in the exploration network. Based on figure 2.6 we argue that the connection of previously unconnected components in 2005 into a large giant component can only take place when organizations form non-local linkages. Given our findings that a giant component has only emerged in the exploration network and not in the exploitation network, we can now say that non-local link formation only plays a role in the exploration of new and relatively un-codified knowledge. This finding is in line with the predictions of Sidhu et al. (2007) who argue that in exploration, non-local linkages might create novel combinations which outweigh the costs of integrating external knowledge.

2.5 Concluding remarks

In this chapter we have provided an empirical overview of how pharmaceutical innovation activities have developed in Europe in the last ten years. In order to do so we have relied on information about alliances which cover the whole production process in drug development. The main actors, such as academic centers, biotech firms and large pharmaceuticals have been identified based on their involvement in these alliances. For these organizations, alliances are an important means to gain access to external knowledge. As previously discussed, different forms of proximities can have an important influence on a firm's ability to access external knowledge. Geographical proximity appears to be important for the location behavior of organizations since more than 70 percent of the organizations in Europe are concentrated in the 30 largest clusters. However, when analyzing the alliance behavior of these clustered organizations in terms of local and non-local search, it appears that geographical proximity or local linkages are of minor importance to these organizations. In chapter 3 the determinants of local and non-local search from a geographical perspective are analyzed in greater detail.

The second part of this chapter has focused on relational proximity as a means for organizations to access external knowledge. Also relational proximity can be expressed in terms of local and non-local search, be it in a network of collaborations. By studying the evolution of the European pharmaceutical network of collaborations over time, we analyzed whether organizations prefer to ally with their existing partners or with partners of their partners, or whether they allied with more distant organizations in the network. We furthermore distinguished whether this relational behavior is different when organizations explore new knowledge or exploit existing knowledge. We found that local search is clearly more prevalent when exploring new knowledge and that with one exception 'distant' knowledge search does not apply to the European scene. This latter finding is explained by the fact that for the most part the European network is highly fragmented. The exception is an interesting development in 2005, when about half of the European pharmaceutical actors are connected to an emerging giant network component. In Chapter 4 we further explore the impact of relational proximity on network evolution.

Chapter 3   Complementarities and the R&D boundaries of the firm: A project level study on pharmaceutical R&D strategies[1]

3.1 Introduction

A firm's ability to innovate is increasingly the result of both internal R&D efforts and external knowledge sourcing (Gambardella, 1992; Freeman, 1991). External knowledge sourcing can be performed through informal personal interactions, formal collaborations, spin-out (and later spin-in) companies and consultancy or through job mobility (Abramovsky et al, 2007). Especially in the biopharmaceutical industry the complementarities between internal R&D and external sourcing through formal collaborations play an important role in large pharmaceutical innovation strategies (Arora & Gambardella, 1990; Henderson & Cockburn, 1996; Pisano, 1990).

Until recently, externally sourcing of R&D has been considered a substitute for in-house R&D activities, i.e. R&D has been perceived as either a make-, or a buy decision of the firm. Exemplifying in this respect is the seminal work of Pisano (1990) on the R&D boundaries of the firm. Although the author acknowledges complementarity between in-house R&D activities to be important, potential complementarities beyond the boundaries of the firm are ignored (Pisano, 1990).

Our study fills an important gap in the literature. While there exists a rich theoretical literature on complementarities between R&D activities, limited data availability has so far constrained empirical testing of some key features of these theories. Using unique data on more than 1300 early stage research projects of large pharmaceutical firms, we empirically study two important issues from the literature.

First, we investigate conditions under which in-house and external research are complementary to one another. Complementarity between two activities is defined as "Adding an activity while the other activity is already performed has a higher

---

[1] This chapter is based on joint work with An Vermeersch and McKinsey and Company.

incremental effect on performance than adding the activity in isolation"[2] (Cassiman & Veugelers, 2006 pp. 70)[3]. While the above definition of complementarity is elegantly simple, it does not provide any direction about when to expect complementarity and how these activities come about to be complementary. We therefore turn to Cohen and Levinthal's notion of absorptive capacity, which is similar to the notion of complementarity. The authors state that firms with a sufficient stock of relevant in-house R&D are better able to achieve complementarities from combining internal with external R&D. Internal know-how is used to effectively screen and absorb external knowledge and to exploit these findings internally (Cohen and Levinthal, 1989). So far, existing empirical work has not provided any answer as to what defines, which knowledge is 'relevant' internal knowledge and how much knowledge is a 'sufficient' stock of knowledge. Our results indicate, which types of pharmaceutical R&D knowledge are relevant to achieve complementarities and, more importantly, we indicate a critical mass of prior R&D that is necessary for complementarities to occur. Second, we investigate in detail how these complementarities occur. According to the theory of absorptive capacity (Cohen & levinthal, 1989; van den Bosch et al, 1999) there is a two-way knowledge flow between internal know-how and external knowledge that underlies the relationship between complementarities and performance. Knowledge flows or spillovers between activities cause learning effects and subsequently increase marginal returns on firm performance[4]. On the one hand, internal know-how is claimed to increase the marginal return to external sourcing through an increased ability to effectively screen and contribute to external projects knowledge (Lane & Lubatkin, 1998). On the other hand, external knowledge that has been absorbed needs to return into the organizations internal knowledge base in order

---

[2] Complementarity between two activities A_1 and A_2 arises only if
$\Pi(1,1) - \Pi(0,1) \geq \Pi(1,-) - \Pi(0,0)$, whereby $\Pi(A_1, A_2)$ represents performance, and each activity A either takes place (1) or does not (0).

[3] Milgrom and Roberts (1990) have first coined the term complementarity to describe synergies among organizational practices within the firm. In the strategy literature the concept of complementarity is better known as 'strategic fit' (Porter, 1980). Strategic fit is defined as: 'the degree to, which the needs, demands, goals, objectives, and/or structures of one component are consistent with the needs, demands, goals, objectives, and/or structures of another component' (Nadler and Tushman, 1980: 36)

[4] Additionally it might be the case that complementarity raises competition (or tournament effects) between activities, which leverages efficiencies and reduces organizational slack. Another potential driver of the relationship between complementarities and performance might be that investment in one activity improves selection capabilities for other activities[4] (Veugelers, 1997).

to truly increase internal innovations. We will refer to the former as knowledge outflow and to the latter as knowledge inflow. We test both knowledge flows and their effect on performance directly. First, we measure how the performance of external R&D projects changes when the firm has generated sufficient in-house R&D in a similar knowledge domain. Second, we measure how internal R&D project performance changes when 'sufficient' external R&D is undertaken in a similar knowledge domain. Our findings are in line with the theory on absorptive capacity. This means that sufficient internal knowledge does not improve the performance of external R&D projects. External R&D projects that are selected do perform exceptionally well, but the main improvement occurs amongst the in-house R&D projects once a few (probable well screened) external projects are added to the 'group' of internal R&D projects. In other words, having a sufficient stock of internal R&D in a relevant knowledge domain enables the firm to screen (attract) the right external R&D projects from which it can learn and subsequently increase its marginal return on internal R&D. Finally, our findings indicate that this knowledge flow (or spillover) only occurs when relatively few external projects are added to the group of in-house R&D. We strongly suspect that the knowledge that is ultimately responsible for realized absorptive capacity does not flow or spill over between projects or even researchers, but only travels because the same people who are working on externally sourced projects are applying the obtained knowledge internally. This would naturally limit the number of external projects in relation to in-house capacity.

3.2 Literature

A number of elementary studies have identified complementarities and some studies have, what might be even more important, identified circumstances that drive complementarity. With our study, we hope to further our current understanding of why and how complementarities occur among research activities within-, and beyond the boundaries of the firm.

The earliest work on complementarities between different research activities goes back to Coase's (1937) work, where he argues that as a firm accumulates R&D experience internally the costs of internalizing new R&D decreases. More specifically, Nelson & Winter (1974) state that the 'ease' of internalizing new R&D depends on whether prior accumulated R&D is similar to the newly acquired R&D,

and that the costs of these activities are reduced because of learning curve effects (Pisano, 1990). Also Milgrom and Roberts (1990) have focused their work on complementarities within the boundaries of the firm, although their definition of complementarity has been applied in some of the more fundamental work on complementarities between internal R&D and externally sourced R&D knowledge as well. The seminal work on complementarities between internal R&D and external sourcing of knowledge is the paper by Cohen and Levinthal (1989) on the two faces of R&D. The authors state that firms invest in R&D primarily to generate internal innovations. More interestingly, they discover that a side-effect of R&D investment is that it enables firms to appropriate external, publicly available spillovers more easily than firms, which invest in R&D to a lesser extend. This side-effect is termed 'Absorptive capacity'. It is defined as a firm's relevant stock of prior knowledge that enables it to identify, assimilate and exploit knowledge from the environment. While the emphasis is on knowledge influx (absorbing external knowledge in), empirical work that followed from this study has often looked at the effect of internal R&D on successful external knowledge sourcing, which implies knowledge outflow (Lane & Lubatkin, 1998; Arora & Gambardella, 1990). The literature review on absorptive capacity by Zahra and George (2002) has brought structure to a growing conceptual ambiguity around absorptive capacity. The authors identify three stages within the construct of absorptive capacity, which are essentially already incorporated in the original definition of Cohen & Levinthal (1989), namely: identification of external knowledge through external sourcing activities. Second, the assimilation or conversion of this knowledge back into the firm and third, the exploitation of the absorbed knowledge to new or improved products and processes. Based on these three phases, most existing work can be grouped as focusing on potential absorptive capacity (phase one and two) or on realized absorptive capacity (phase three) (Zahra & George, 2002). An interesting study that does cover all three phases of absorptive capacity (identification, incorporation and exploitation) is the study by Cassiman & Veugelers (2006). The authors find that firms, which 'make and buy' R&D have a higher marginal return on innovation than firms who only 'make' or only 'buy', especially if they are more heavily involved in basic R&D. More precisely, a 10% increase in reliance on basic R&D increases the likelihood of combining internal and external sourcing by 2.7 % (Cassiman & Veugelers, 2006 pp. 77). The studies described above have convincingly argued that complementarities exist among R&D

activities within and between firm boundaries. An important condition for complementarities to occur, as Cohen & Levinthal already mentioned, is that internal and external knowledge are relevant to one another. With two exceptions, existing contributions have largely ignored this condition of knowledge relevance. One exception is the work of Arora and Gambardella (1994) who test whether firms use their external linkages as complements. While complementarity itself is not precisely defined in this study, it is assumed that different types of external linkages are complements if they do not have overlapping (knowledge) purposes. Moreover, the authors argue that external linkages such as research agreements with universities and acquisitions of biotech firms are complementary strategies because they serve different purposes but are still correlated. While indeed, it is generally acknowledged that activities, which are completely overlapping in terms of purpose or knowledge domain are considered as substitutes, it is not very clear whether non-overlapping features defines them as complements (Besanko, 2007). A more precise investigation into the relevance of knowledge between R&D activities is provided in the work of Lane and Lubatkin (1998). These authors argue that a firm's absorptive capacity is often seen as a firm-specific characteristic that determines its innovativeness vis-à-vis others to a large extent. However, Lane and Lubatkin (1998) state that a firm's ability to absorb external knowledge depends on the 'type' of external knowledge (and the partner carrying this knowledge) the firm is absorbing. In other words, absorptive capacity is a relational characteristic rather that an actor characteristic of the firm, since it differs with each external partner. A firm's absorptive capacity in this sense depends on the knowledge (cognitive) similarity between internal knowledge (experience), and external knowledge, and thereby becomes a relational characteristic of the firm. This 'new' notion of absorptive capacity furthermore implies that innovativeness, which is increased by absorptive capacity, differs for each activity where a firm is tapping into a new external source of knowledge. An important implication of this finding is that measuring absorptive capacity requires project-level information.

Our study differs and enriches these existing studies in a number of ways. To start with, all above mentioned studies are performed at the level of the firm, while our study enables a direct measure of performance at project level. Having project level information has important advantages. Not the least advantage is that it allows us to circumvent the danger of firm heterogeneity driving endogenous decisions of which

projects are selected[5]. Furthermore, while being specific for the pharmaceutical industry in which our research is situated, we identify the type of knowledge where learning curve effects and subsequent complementarities occur. Most importantly however is that while previous studies have convincingly shown the existence of complementarity as a binary choice, our study allows us to treat complementarity as a continuous variable. More specifically, we identify a size threshold over which complementarities occur, which concretizes Cohen & Levinthal's (1989) notion of the 'sufficient' stock of knowledge required for achieving complementarities. Given the existence of complementarity, our data allow us to empirically disentangle whether these complementarities represent potential absorptive capacity (arising from improved external sourcing) or whether the firm has indeed managed to reintegrate and exploit external knowledge, which represents realized absorptive capacity.

---

[5] For additional information on how to deal with the problem of unobserved firm heterogeneity we refer to Cassiman & Veugelers (2006).

3.3 Empirical setting: the pharmaceutical industry

A number of developments in the pharmaceutical industry over the last decade have made the quest for a strategy that achieves complementarities and subsequent innovation capacity an important and largely unanswered question in this industry. Within big Pharma there is an increased pressure on R&D due to decreased R&D productivity and approaching patent expirations. Along the pressure for big pharma to maximize shareholder value, strategies have shifted in the last two decades from being research driven to being market driven (Drews, 2003). In a research driven environment, emphasis of decision making was within R&D departments, where managers were geared at innovations originating from deep knowledge of disease pathways and pharmacology. While the industry was consolidating in the beginning of the nineties, the gravity of decision making has shifted toward marketing and finance departments (Drews, 2003). This strategic shift implied a more quantity based approach toward research whereby drug discovery was increasingly considered a statistical event (Booth & Zemmel, 2004). Today, it appears that this quantitative approach toward research is not yet paying off in terms of productivity, and analysts are revaluating the early days' in vivo empiricism based on disease knowledge (Erickson, 2003).

Furthermore, a series of findings suggest that alternatives to the traditional in-house R&D model of big pharma might be more successful. The first finding concerns the higher success probabilities of (new) biotechnologies. The proportion of newly admitted compounds using biotechnologies has increased with 20 -25% compared to the more traditional chemical based compounds (Reichert, 2000). Biotechnologies are mainly exploited by small-, and medium sized biotechnology firms, while chemical based compounds mainly originate from big pharma. Second, newly developed compounds that are discovered in-house are being outperformed by compounds that are produced externally or through external collaborations[6]. Moreover, as biotechnologies increase the scope of research, big pharma increasingly realize that it is impossible to cover the whole spectrum of technologies themselves. While research for new drugs was traditionally conducted in-house, large pharmaceutical firms have by now build research portfolios where internal R&D efforts are combined with

---

[6] 'Improving the pharma research pipeline' *McKinsey Quarterly*, August 2004.

external R&D collaborations. Some analysts even go as far as to state that Pharma companies should be virtual in research, which means that they in-license all compounds from preclinical testing onwards. Against the background of these developments it has become questionable whether big pharma still has to play a role in research. Wouldn't it be better if big pharma narrows down its core competences to downstream drug development and employ a virtual research model? Or are there still advantages to be obtained from in-house research? If the latter is true, which portfolio would generate complementarities between internal and externally sourced research projects?

3.4 Empirical strategy

Our empirical strategy is summarized in figure 3.1.

Fig. 3.1 - Empirical strategy

| Select big pharma companies | Built research project database | Define project characteristics | Analyze relation between project characteristics and success prob. |
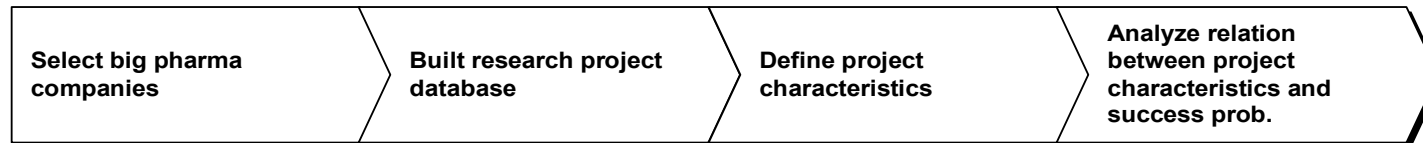
- Select top 20 pharma companies (source: Evaluate) based on highest R&D spending in 2005

- Build research portfolio 2000-2002*** (source: PharmaProjects): 1328 newly announced **research projects** in preclinical stage** between January 2000 and December 2002
- Define dependent variable: success probability
  - Success if project reaches clinical Phase I by Jan 1, 2007
  - Failure if project cancelled or no info for ≥4 years
  - Unknown if no info for 1-3 years

- Allocate project characteristics to each project
  - Success vs failure (reaching clinical phase) (SUCCESS)
  - Internal vs. External sourced (EXTERNAL)
  - Biological vs. chemical (BIOTECH)
  - Formulation versus new drug (FORMULATION)
  - Disease area (also therapeutic area, target, pharmacological activity) (e.g CANCER, INFECTION)
- Allocate grouped project characteristics to each project:
  - Size of group* (SIZE)
  - Share of external projects in group (FEWEXTERNALS)

- Use Binary logistic regression analysis define which group-, and project characteristics have higher probability of success

* Number of projects that belong to same disease area (DASIZE), target (TARGETSIZE), therapeutic area (TASIZE) or pharmacological (PHARMACOSIZE) activity

** All stages of preclinical investigation including discovery, research, lead optimization.

*** Following Phelps (2003)5 we assume a project duration of three years. As a result, all projects that started between 2000 and 2002 make up the 2002 research portfolio.

3.4.1 Data selection

We selected the 20 highest R&D spending pharmaceutical firms in 2005 from Evaluate Pharma, a frequently used database in the pharmaceutical industry for forecasting and analysis services. Of the selected firms, we used Pharma Projects database to extract the whole research portfolio per firm at a certain time. PharmaProjects is a privately held project monitoring firm, which continuously searches for information on both internal and externally sourced projects of large pharmaceutical firms through a number of search channels. Primarily, PharmaProjects visits events and conferences where pharmaceutical firms meet to exchange information about the projects running through their pipelines. This information is then verified and updated with press releases, website information and annual reports. Telephone surveys are regularly conducted to verify the accuracy of their database. As a deliberate strategy, no use is made of patent information, since "often..." our informant claims, "...the firm files patents on anything that lies around in the lab to create a smoke screen and hide their actual R&D strategy" (information based on telephone interview with Pharmaprojects data manager).

We focus our analysis on research projects that are active at the earliest stage of research before entering clinical testing[7]. In doing so, we follow earlier work by Pisano (1990) who convincingly argues that from clinical testing onwards external contracting is often done by technological licensing instead of R&D contracting. Technological licensing and R&D contracting are two fundamentally different contracts. R&D contracts are typically long-term agreements where knowledge exchange and learning effects take place, while technological licensing agreements are one-time exchanges to obtain rights to an already developed technology (Pisano, 1990 pp. 163). As in our study we are interested in complementarities arising from learning through knowledge spillovers between and within organizations, we restrict our analysis to these early stage research projects. Another reason for this restriction is that our dependent variable (probability to enter clinical I) is more reliable when restricted to early stage research. If we were to measure the probability of a project to reach the market for example, we would be unable to distinguish complementarities

---

[7] Projects only entered the database if they had a solid chemical structure and a therapeutic goal had been identified.

from many other factors affecting whether a projects survives the 12 year (or longer) ride through the pharmaceutical production process.

3.4.2 Building a research portfolio

To determine each firm's research portfolio of 2002, we summed all research project that were announced as annual newly entering projects in early research from 2000 until 2002. We thereby follow a study by Phelps (2003) who shows that the average duration of R&D projects is three years (see also: Phlippen & van der Knaap, 2007). As a result, all newly announced projects in 2000, 2001 and 2002 are assumed to be part of a firm's 2002 research portfolio. Our choice of constructing the 2002 research portfolio and not a more recent portfolio is related to our dependent variable 'probability to reach clinical testing'. A project in early research can take (on average) up to four years to reach clinical testing, which makes it necessary to track each project until the end of 2006 to know whether it has been successful in reaching clinical testing on humans. Our initial sample consisted of 1328 early stage research projects (before clinical testing I). Leaving out projects of which no success probability was known reduced our sample to 977 projects. Furthermore, we excluded all projects where no disease area information was given, which reduced our final work set to 762 projects.

3.4.3 Variables

Dependent variable: success probability

Each project is defined as either successful, failure or unknown in reaching the first stage of clinical testing (SUCCESS). If we didn't find the project in our database for 4 years or longer, we decided to label it as a failure. Projects were labeled unknown if there was no information between 1 and 3 years. We used Binary logistic regression analysis to model the success probabilities of a research project.

Control variables

In every regression we controlled for a number of variables that are strongly associated with the probability of a project reaching clinical testing successfully (see: empirical setting). The first is the variable indicating whether a project is being developed in-house or through external collaboration (EXTERNAL), as previous work has shown external collaborations to perform better on average than internal projects over the whole production process[8]. A second control variable indicates whether a project involves chemicals or biotechnologies (BIOTECH), as biotechnology based compounds are found to outcompete compounds based on chemical substance. We further controlled for the effect of a project aiming at a reformulation of an existing drug or aiming at a new drug (FORMULATION), since the former are assumed to be more likely to reach clinical testing. Each project is focused at a certain disease area. To prevent the disease area (e.g. anticancer or inflammation) itself to be driving the success probabilities of our projects we included dummies for all disease areas in most of our regressions (ANTICAN, INFLAM etc). At the level of the firm, we control for firm heterogeneity simply by adding dummies for all firms in most of our regressions (ROCHE, GSK, etc).

Explanatory variables

There are two main sets of explanatory variables. The first set is aimed at identifying the conditions for complementarity to arise, and the second set is aimed at understanding how complementarities occur between internal projects and external projects, i.e. whether they result from potential (knowledge outflow) or from realized absorptive capacity (knowledge inflow).

Conditions for complementarity

As we have argued in our introduction, Cohen & Levinthal (1989) and the work thereafter has emphasized the importance of a firm's 'sufficient stock of relevant prior knowledge' in order to absorb external knowledge effectively. We test what constitutes a 'sufficient' stock and what knowledge is 'relevant' to obtain complementarities.

---

[8] 'Improving the pharma research pipeline'. *McKinsey Quarterly*, August 2004

Where to look for complementarities?

Which type of knowledge similarity generates complementarities? While it is typically assumed that projects aiming at similar therapeutical areas might compete and/or learn from each other and create complementarities by doing so, we test 3 other potential types of knowledge areas where complementarities might occur. First we add disease area as a potential knowledge area that might be an alternative for therapeutic area. TAsize (therapeutic area) is a more broadly defined group than DAsize (disease area). For example the therapeutical area named 'cognition enhancer' consists among others of the disease area 'Alzheimer'. Another knowledge area where complementarities might occur is the target that a drug in a project is aiming at (TARGETsize). A drug target can be the protein to which the drug binds, inhibits or activates (e.g. receptor subunits or enzymes). Finally we tested whether the pharmacological activity, which describes the beneficial or adverse effects of a drug on living matter (i.e. it describes how the drug works) might generate complementarities (PHARMACOsize). Based on each of these four knowledge areas we grouped all projects within each firm and tested whether more projects in each of these knowledge areas (within a firm) increases the average performance of projects.

When is a stock of knowledge sufficiently large?

The aim is to determine the effect of the number of 'similar' projects in a firm's research portfolio (SIZE) on the probability of a project reaching clinical testing. The knowledge area to which 'similarity' applies is to be defined as a first step in our analysis. This variable (SIZE) is categorized as either 'no or just one similar project' (SMALL), 2-9 similar projects (MEDIUM), or containing 10 or more similar projects (LARGE)[9]. Complementarities arise when more similar projects leads to higher average success probabilities. We thus expect that projects in category LARGE perform better than projects in category MEDIUM. Projects categorized as SMALL contain (nearly) isolated projects, which by definition do not measure

---

[9] We explored different size categories and the currently used categories were most able to discriminate amongst success probabilities.

complementarity. However, being a relatively large group of projects, we used this category as our reference category to benchmark our two other categories against.

How do complementarities arise between internal and external projects?

The theory of absorptive capacity provides guidance on how complementarities are expected to arise between internal research and externally sourced research projects. While a firm's absorptive capacity implies knowledge flowing outside in, this theory argues that internal knowledge is first used to screen and absorb external knowledge, and as a second step the firm reintegrates the absorbed knowledge internally to improve internal R&D. We measure both parts of this process separately by identifying:

1. How the performance of external projects changes when a 'sufficient' amount of 'similar' internal projects are running (EXTERN_SUCCESS).

2. How the performance of internal projects changes when a 'sufficient' amount of 'similar' external projects are running (INTERN_SUCCESS).

The question about 'sufficient' amounts of internal and externally sourced projects essentially asks how a firm should design its research portfolio with regard to the ratio of internal and external R&D investment. Surprisingly we found no study that deals with this question explicitly and hence we explored the 'optimal' external/ internal ratio among 'similar' projects ourselves.

3.5. Results

3.5.1 Effect of individual project characteristics

Our analysis starts with the assessment of individual project characteristics that have in previous studies been identified as having a significant impact on the success probabilities of R&D. The two main characteristics are whether a project is conducted in-house or through external sourcing, and second whether a project builds on chemicals or on biotechnologies[10] [11].

Fig. 3.2 – Individual project characteristics
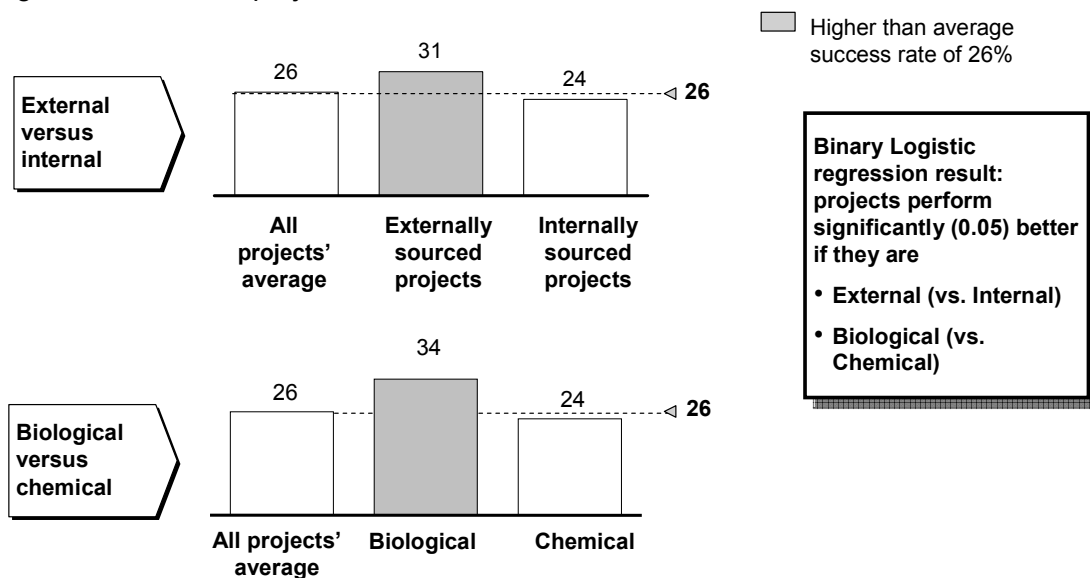
Externally sourced projects are significantly more likely to reach clinical testing (our indicator of success) than internal projects. While previous studies have already shown this differences between internal and externally sourced projects at the development stages in R&D (clinical I to III) (DiMasi, 2001), our findings indicate that the advantage of external collaboration already occurs during early research. We further found that projects involving biotechnologies, such as recombinant DNA technologies or monoclonal antibodies are also significantly more likely to be successful compared to projects based on chemical compounds. Figure 3.2 illustrates

---

[10] See chapter 1 section 1.3 for an overview.
[11] We also tested the effect of a project being a new formulation of an existing drug or a 'real' new drug.

these findings by showing the differences in average success rates related to external collaborations and biotechnologies.

3.5.2 Conditions for complementarities

Identifying 'relevant' knowledge

Our data allowed us to test four knowledge areas where complementarities might occur. More specifically, we grouped projects around the same disease area, the same target, the same pharmacological activity and around the same therapeutic area (within a company). For each grouping we then analyzed the effect on success probabilities, as is shown in table 3.1. After controlling for individual project characteristics we found that only disease area grouping has a positive significant impact on success probabilities, and that neither grouping by target, by pharmacological activity or by therapeutic area has a significant impact on success probabilities[12]. From here on, our analysis focuses on projects grouped around disease areas, as it appears to be the only relevant knowledge area for achieving complementarities.

---

[12] Interestingly, grouping based on therapeutic area appears to have a significant negative effect on success probabilities. It might be the case that on the higher aggregation level that therapeutical area represent, competition outweighs learning effects.

Table 3.1 – Identifying relevant knowledge areas.

Effect of grouping by disease area, target, pharmacological activity and therapeutical area on a projects' success probability

		B	S.E.	Sig.	Exp(B)
Step 1(a)	BIOTECH	.724***	.234	.002	2.064
	FORMULATION	1.129***	.378	.003	3.092
	EXTERNAL	.360*	.195	.065	1.434
	Ta size	-.049***	.016	.003	.952
	PharmacoclusterSize	-.015	.011	.168	.985
	Da size	.034**	.017	.044	1.035
	TARGETsize	.089	.071	.209	1.093
	Da externals	-.082	.050	.102	.921
	Constant	-1.008	.164	.000	.365

a Variable(s) entered on step 1: BIOTECH, FORMULATION, EXTERNAL, TAsize, PharmacoclusterSize, DAsize, TARGETsize, DA Externals

Notes: coefficients significant at 1%***, 5%**, 10%*

Identifying a sufficiently large stock of knowledge

Previous work on complementarities has provided clear evidence that adding an R&D project to an already existing stock of R&D has a higher incremental effect on innovation than adding an R&D project in isolation (Milgrom & Roberts, 1990). Going one step further, we analyze whether the size of the existing stock of knowledge matters for achieving these complementarities. Intuitively, one could imagine that a critical mass of relevant knowledge must be achieved in order to truly benefit from complementarities. In order to test whether this is indeed the case we categorized projects as either belonging to a large disease area group, (i.e. containing 10 projects or more per firm), to a medium sized disease area group (containing 2 to 9 projects within the same disease area per firm), or as focusing on a (nearly) isolated disease area. In the latter category success probabilities are not caused by complementarities. Based on the actual distribution of projects over disease areas by our set of firms in figure 3.3 we chose the boundary between large and medium sized groups.

Fig. 3.3 – Number of projects per disease area.
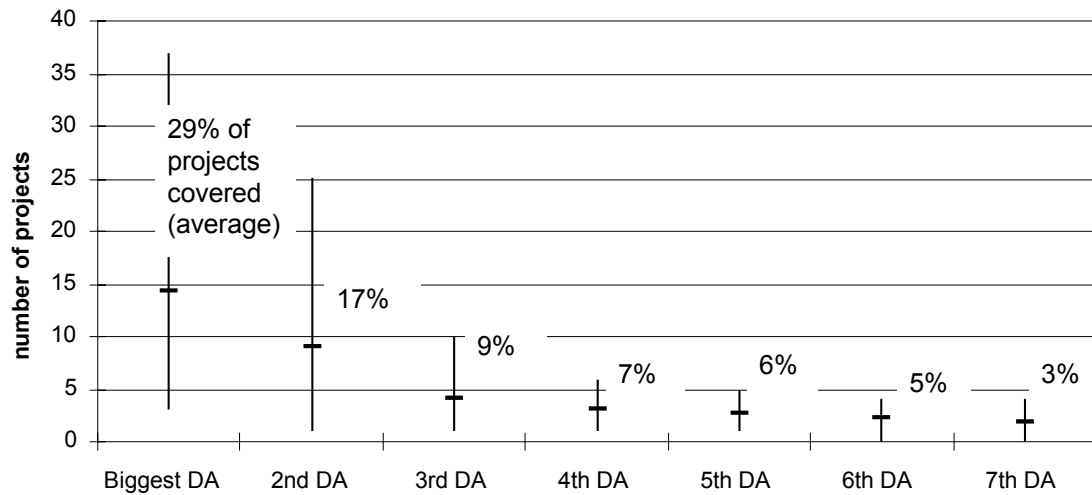
Figure 3.3 reveals that the largest R&D spending pharmaceutical firms invest 46 percent of their total research projects in disease areas with on average 10 projects or more (within the 2 largest disease areas). If one can speak of a critical mass of knowledge that drives complementarities, we expect the boundaries to be around 10 or more projects[13].

Table 3.2 shows the binomial Logit regression results on the effect of the group size to which a project belongs on the probability that the project reaches clinical testing.

---

[13] We also explored shifting the boundaries to 9 or more projects and to 12 or more projects.

Table 3.2. – Identifying a sufficient stock of knowledge.

Effect of the group size to which a project belongs to on its success probability

		B	S.E.	Sig.	Exp(B)
Step 1(a)	BIOTECH	.524	.198	.008	1.689
	FORMULATION	1.227	.363	.001	3.409
	EXTERNAL	.375	.164	.023	1.454
	LARGE	.367	.211	.082	1.444
	SMALL	1.020	.181	.000	2.773
	INFECTION	.461	.230	.045	1.585
	Constant	-1.815	.145	.000	.163
Step 2(b)	BIOTECH	.518***	.199	.009	1.679
	FORMULATION	1.229***	.363	.001	3.417
	EXTERNAL	.375**	.165	.023	1.456
	LARGE	.371*	.211	.079	1.448
	SMALL	1.035***	.182	.000	2.814
	Astrazeneca	.638**	.318	.045	1.893
	INFECTION	.495**	.230	.032	1.640
	Constant	-1.862	.148	.00	.155

a Variable(s) entered on step 1: INFECTION. b Variable(s) entered on step 2: Astrazeneca.
Notes: coefficients significant at 1%***, 5%**, 10%* Notes2: N = 977.

In order to exclude any effect of either firm heterogeneity or effects that are specific for any disease area, we added firm dummies and disease area dummies next to the usual project characteristics as control variables. To keep our results readable, we used a stepwise selection method for our variables with entry testing based on the significance of the score statistic, and removal testing based on the probability of a likelihood-ratio statistic based on the maximum partial likelihood estimates. Our main variable of interest is LARGE, which is offset against our base variable MEDIUM. As table 3.2 reveals, the variable LARGE is significant at the 10 percent level, which indicates that projects active in a disease area where at least 10 other projects are running are more likely to reach clinical testing compared to projects in a disease area where only 2 to 9 similar projects are running. This finding confirms that projects in which a firm has built a critical mass of disease knowledge have a higher marginal return to performance than projects in which a firm has not built such a critical mass.

More generally, complementarities between research projects arise once a firm has built a critical mass of knowledge in a similar knowledge domain.

However, one could argue that there is a critical danger of endogeneity driving our results i.e., projects do not become more successful because of complementarity effects, but the management of a firm puts its 'golden eggs' in one basket, namely in its main disease areas. This would imply that the expected success of certain projects drives management to focus their attention on these projects and create a large number of similar projects around these 'golden eggs'. To test whether this is the case, we analyze the effect of each firm's two main disease areas[14] on the success probabilities of projects in these DA's. Table 3.3 shows the results of this test, whereby we replicated the test on size effects (table 3.2) while interchanging the categories LARGE, MEDIUM, SMALL with the categories DA2largest and OTHER (reference category) [15].

---

[14] Again, we chose to analyze a firms two main DA's since these represent 46 % of all research projects, while the third largest disease areas and beyond are strongly decreasing their contribution to the firms research portfolio (see figure 3).

[15] To prevent our main variable of interest DA2largest to be excluded from the results based on restricted entry testing used in table 2, we unconditionally let the variable DA2largest and the BIOTECH, FORMULATION and EXTERNAL variables enter our equation.

Table 3.3 – Effect of projects belonging to a firm's 2 largest disease areas on success probabilities

		B	S.E.	Sig.	Exp(B)
Step 1(a)	BIOTECH	.510	.221	.021	1.665
	FORMULATION	1.226	.374	.001	3.407
	EXTERNAL	.243	.183	.183	1.276
	DA2largest	-.311	.170	.068	.733
	Wyeth	.703	.336	.036	2.020
	Constant	-1.046	.124	.000	.351
Step 2(b)	BIOTECH	.510**	.221	.021	1.665
	FORMULATION	1.245***	.375	.001	3.472
	EXTERNAL	.237	.183	.196	1.268
	DA2largest	-.333**	.171	.052	.717
	GSK	.548**	.278	.049	1.730
	Wyeth	.759**	.338	.025	2.136
	Constant	-1.089	.126	.000	.336

a Variable(s) entered on step 1: Wyeth. b Variable(s) entered on step 2: GSK.
Notes: coefficients significant at 1%***, 5%**, 10%*
Notes2: N = 977.

DA2largest measures whether a project is part of a firms two main disease areas or not. As becomes clear in table 3.3 projects that belong to one of the firm's two most important disease areas does not increase their probability of reaching clinical testing. Moreover, there seems to be a negative effect from being part of a firm's 2 main disease areas. This effect might be caused by the fact that the firms in our sample differ with respect to their R&D investment strategy. While some firms choose (or are able) to built a critical mass in a few disease areas, other firms rather spread their R&D projects over different disease areas. These strategy differences become clear when looking at figure 3.3. The number of projects that form a firm's 2 largest disease areas range from 2 to 37 projects. Interestingly, the strategy of spreading projects over different disease areas, referred to as risk diversification strategies, is appearing to be paying off as well. Although this study investigates complementarity effects among projects, the highly significant positive effect of nearly isolated projects (SMALL) in table 3.2 raises our suspicion that aiming for complementarities by building a critical mass of disease knowledge is not the only rewarding strategy.

3.5.3 How do complementarities arise between internal and external projects?

So far, we have found that complementarities arise from grouping a relatively large amount of research projects around a disease area. As a next step, we focus on these large disease area groups to find out how the ratio of internal R&D and external R&D affects complementarities among internal projects and externally sourced projects separately[16]. The ratio between internal and external R&D has been surprisingly little discussed within the literature on absorptive capacity. We argue however that it is a fundamental question since the capacity to absorb larger amounts of external information must depend greatly on a larger internal capacity to absorb this information. To some extend the work of Lane and Lubatkin (1998) recognizes the importance of this ratio by focusing on relative absorptive capacity. However, they do so at the level of the dyad (i.e. a relation between internal R&D and one externally sourced project).

Once the ratio of internal versus external projects that generates complementarities is determined, we can test whether external projects or internal projects are most responsible for the increased success probabilities. In this part of the analysis we use a subsample of our data, namely only projects belonging to a firm's large disease area (10 or more similar projects). Figure 3.4 below plots the number of projects that firms run in large disease areas at different internal / external ratios.

---

[16] By reducing our sample to projects in large disease areas we are unable to control for firm-, and disease fixed effects. This is due to the fact that only few firms are able to create a large number of projects in a few disease areas (e.g. anticancer and infection).

Fig. 3.4 – The number of projects in large disease areas (within the firm) at different external/internal ratios
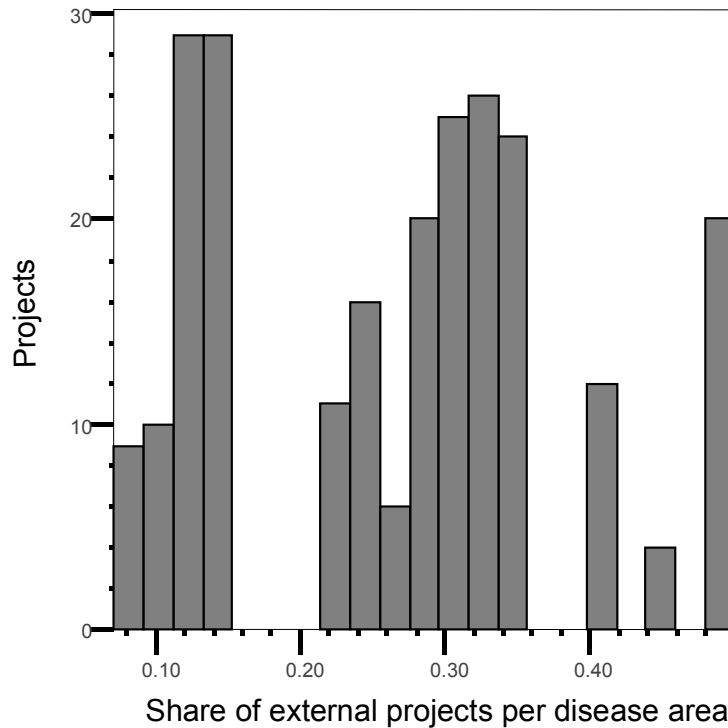
Based on the above plot we divided projects as either belonging to a disease area with few external projects and many internal projects (ratio externals 20/80 or less) or vice versa (ratio externals 20/80 or more)[17]. Dividing our projects into these to categories allows us to test if projects perform differently in each category. If they do, we can test whether internal projects benefit from a specific internal/external ratio or whether external projects benefit from this ratio. Projects in the 'poor performing' category serve as our benchmark. The results of these tests are displayed in table 3.4.

---

[17] Obviously this choice is somewhat arbitrary. We explored different ratios and found this division to be discriminating our success probabilities.

Table 3.4 - Effect of (internal) projects belonging to a large disease area with less than 20% external R&D projects on success probability.

	B	S.E.	Sig.	Exp(B)
BIOTECH	1.416***	.508	.005	4.120
FORMULATION	-20.317	40192.970	1.000	.000
EXTERNAL	1.022*	.561	.069	2.778
FEWEXTERNALS	1.040**	.493	.035	2.829
EXTERNAL by FEWEXTERNALS	-1.071	1.062	.313	.343
Constant	-1.907	.391	.000	.148

a Variable's) entered on step 1: BIOTECH, FORMULATION, EXTERNAL, FEWEXTERNALS, EXTERNAL FEWEXTERNALS .
Notes: coefficients significant at 1%***, 5%**, 10%*
Notes2: N = 136

In table 3.4 the variable FEWEXTERNALS shows that projects in a disease area with an external ratio of 20/80 or less perform significantly better than projects in a disease area with a higher externals ratio. While this variable does not distinguish among internal and external projects, the addition of our interaction term (FEWEXTERNALS by EXTERNALS) controls for the (slightly negative) effect that external R&D projects have in the disease areas with a low external ratio. To put it differently, the higher performance of projects in DAs with low externals ratio is mainly attributable to the improved performance of internal projects. To clarify this point we add a fifth table where we only consider the subsample of internal projects in large disease areas.

Table 3.5 - Internal project performance in large clusters with few externals compared to other internal projects.

	B	S.E.	Sig.	Exp(B)
Step 1(a) BIOTECH	.923**	.418	.027	2.516
FORMULATION	.380	1.182	.748	1.463
INTERNAL_SUCCESS	.581*	.353	.100	1.788
Constant	-1.479	.254	.000	.228

a Variable(s) entered on step 1: BIOTECH, FORMULATION, INTERNAL_SUCCESS.
Notes: coefficients significant at 1%***, 5%**, 10%*
Notes2: N = 182

Here, in table 3.5 the variable (INTERNAL_SUCCESS) indicates the performance of internal projects in large disease areas with a low external ratio in comparison to other internal projects in large disease areas. The results confirm the findings displayed in table 3.4 that internal projects perform better if they are part of a large disease area with few externally sourced projects involved.

3.6 Conclusions

For more than a decade firms in research driven environments such as the pharmaceutical industry, experience an increased pressure on R&D due to decreased R&D productivity and approaching patent expirations (Drews, 2001). In response to this, firms are exploring alternative ways to organize their R&D portfolio. While traditionally early stage R&D has been conducted mainly inside the firm's own R&D laboratories, the last decade has brought forward a huge increase in R&D collaborations with market based firms at all stages of the drug development process. This raises the question whether the make-or-buy decision of the firm should be replaced by a make-and-buy decision of the firm. The answer depends crucially on the extent towhich internal R&D (make) and externally sourced R&D (buy) can be

complementary, i.e. whether performing internal R&D in combination with external R&D generates higher marginal performance than only internal or only external R&D. In theory, complementarity between internal and external R&D exists if a firm has built a sufficient stock of relevant internal knowledge to effectively absorb external knowledge (Cohen & Levinthal, 1989). Our study is the first empirical work that has investigated the conditions for complementarity to arise, and the process through which it occurs. More specifically, we examined what 'type' of knowledge is relevant for complementarity, how much of this knowledge is sufficient, and lastly we examined how this knowledge flows between internal and externally sourced R&D. As it turns out, large pharmaceutical firms can achieve higher marginal returns on their R&D projects, if they group a relatively large number of projects (more than 10) around a specific disease area. By focusing on a specific disease area firms can develop deep in-house expertise, attract the best talent, and be a preferred partner for deal-opportunities outside the firm. Moreover, we conclude that if no more than 20 percent of these projects are externally sourced, complementarity effects are highest. This is caused by the fact that internal projects perform better when the number of externally sourced projects is relatively low. The reason for this might be that knowledge can only be transferred from external projects to internal projects if the same expert scientists are involved in both internal and external partners work. This would naturally limit the number of externally sourced projects. This interpretation is in line with the notion that knowledge required for pharmaceutical drug discovery is highly tacit and embedded in the scientists involved, which makes transfer of this knowledge between people leave alone projects difficult.

Chapter 4   When clusters become networks[1]:
Alliance formation in regional clusters[2]

4.1 Introduction

Regional innovation, local knowledge spill-over and cluster synergy are concepts that have long spurred policy makers to invest heavily in the co-location of firms such as technology- and science parks. These investments however rely on two assumptions: one is that there is some sort of knowledge diffusion process going on among co-located organizations that is beneficial to these organizations, and the other is that this knowledge diffusion occurs in geographically confined areas. These assumptions are supported by empirical and theoretical evidence where high-tech clusters with strongly connected organizations are recognized as engines of national economic performance (Storper, 1995; Scott, 1993; Saxenian, 1994). While there are examples of highly innovative regions where firms exchange knowledge intensively, there are many more regions where co-location does not induce any knowledge exchange.

As the exchange of knowledge is considered crucial for innovation in science based industries, this chapter aims to bring to light what determines knowledge exchange in clusters of co-located firms. More specifically, we exploit empirical data from the European pharmaceutical industry, as it is the most science based industry today and we focus on one particular form of knowledge exchange, namely formal collaborations.

Knowledge can diffuse through various routes, depending among others on the type of knowledge that is diffused (e.g. whether it concerns tacit or codified information or whether it is appropriable) and on the type or organizations involved. Knowledge diffusion can take the form of informal personal interactions, formal collaborations, spin-out companies and consultancy or through job mobility (Abramovsky et al, 2007). In this study we look at formal collaborations as a means of knowledge diffusion. Firms in technology intensive environments transfer valuable knowledge through formal interactions rather than through informal social contacts (Zaheer and

George, 2004). Moreover, we argue that while informal knowledge exchange can play an important role in innovation, it often travels along more formal interactions such as lunch meetings, or other social occasions with formal alliance partners. Especially in our empirical setting, which is the pharmaceutical industry formal collaborations have shown to be crucial in the organization of innovative labor and in the acquisition of new skills and technologies (Hagedoorn, 2002; Owen-Smith et al, 2002).

Given the importance of formal collaborations, the question is: when do these collaborations require geographical proximity? While there is convincing evidence of the existence of geographically mediated knowledge diffusion in science driven industries (see Audretch & Feldman, 1996; Jaffe, 1989; Prevezer, 1997), less is known about what drives this geographically mediated knowledge diffusion or, as in our case, local collaboration. From a policy maker's perspective, for whom geographical proximity is given, one should turn this question around by asking: when do co-located organizations collaborate?

In the existing literature on firm innovation in relation to locational decision, emphasis is mainly on why firms cluster geographically. One important finding is that geographically mediated knowledge spillovers can partly explain the geographical clustering of innovative firms. While studies emanating from US data indeed show that innovative tacit knowledge transfers locally through formal alliances (Zucker et al, 1996), little is known about the European situation. Our study not only reveals partly contradicting evidence, but also indicates factors that do induce local link formation in European clusters. Proposed explanatory factors from the literature that we include in our analysis are: type of organizations present, relational embeddedness, nature of knowledge that is exchanged, technological diversity, and life-cycle effects.

The empirical setting of this chapter is the European biopharmaceutical industry. The biopharmaceutical industry appears to be a very appropriate setting to test our research questions since it is not only a highly innovative science based industry, but firms active in biopharmaceuticals tend to cluster geographically (Swann & Prevezer, 1996; Zaheer & George, 2004; Zucker, 1996). Thereby, due to the specific characteristics of drug development, such as high commercial values and natural excludability, the number of strategic collaborations is high (Hagedoorn, 2002). With rare exceptions, existing studies on innovative clusters are case studies of one or two

regional clusters and existing quantitative studies are mostly based on data from the United States. This chapter distinguishes itself from these existing studies by covering 100 European clusters of deal active firms and fairly detailed information per cluster, such as type of organizations, type of alliances (R&D collaboration, licensing etc) and the therapeutic focus of alliances. Furthermore the time span of our data (1996-2005) allows us to first provide a preliminary longitudinal view of collaboration activity in Europe and finally to create clusters of firms in which alliance activity is based on realistic assumptions of alliance duration. Lastly, our use of longitudinal data enables us to significantly reduce endogeneity problems. In what follows we propose hypothesis based on our discussion of the literature on innovation, collaboration and geographical proximity (section 3). Section 4 of the chapter will discuss the data and research methodology, followed by the results in section 5. In the concluding section (section 6) of the chapter we will discuss the results, its implications and limitations and give some directions for future research. We will start with a preliminary view on the data (section 2).

4.2 Preliminary data view

In figure 4.1 we have plotted the amount of annual newly announced agreements between organizations in the European pharmaceutical industry. We distinguish between local agreements and non-local agreements. Local agreements are agreements between organizations located in the same geographical cluster and non-local agreements are agreements between organizations that are not located in the same cluster. We further define a cluster as a group of co-located organizations. The boundaries of our clusters are chosen in such a way that the geographical distance between organizations *within* our clusters is minimized while maximizing the distance *between* clusters. More details about the boundary setting of our clusters are given in section 4. Agreements, hereafter referred to as linkages, are formed at all phases in the production process of drug development, ranging from early stage drug discovery collaborations to marketing and distribution agreements.

Fig. 4.1 - Annual new collaborative non-local agreements (dotted line) and local agreements (full line)
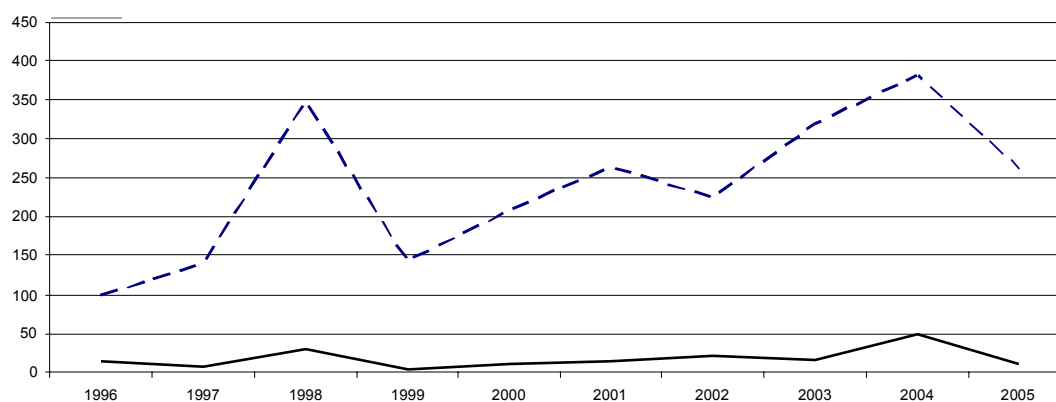
From 1996 till 2005 we can clearly distinguish 3 waves of link formation, with peaks in 1998, 2001 and 2004 respectively. While local linkages and non-local linkages seem to follow the same pattern in link formation, the absolute number of linkages that occur locally is much lower compared to non-local linkages. To identify the whole network of collaborations at various moments in time we created 'snapshots' of the network of collaborations in June 1999, June 2002 and in June 2005. These 'snapshots' build on the assumption of alliance duration of 3 years (Phelps, 2003),, which means that each 'snapshot' captures all new agreements announced back to

three years before taking the 'snapshot'. In table 4.1 and figures 4.2, and 4.3 we show these snapshots of the European pharmaceutical industry. We are aware that our dataset is limited in the sense that we do not observe firms, which have not formed alliances for three years. However, we do not consider this a major problem since competitive firms in the pharmaceutical industry are almost always deal-active (Arora & Gambardella, 1990).

Table 4.1 – European pharmaceutical clusters (descriptive)

	June 1999	June 2002	June 2005
Mean number of organizations per cluster	5.79	7.32	10.04
Standard dev of organizations per cluster	4.96	7.73	11.10
Linkages	687	715	1157
Organizations	303	411	609

European pharmaceutical clusters are increasing in size with mean cluster sizes almost doubling from 1999 until 2005. In the same period we see a growing dominance of London in figure 4.2 as the largest pharmaceutical cluster in terms of deal-active organizations present. This visual observation is confirmed as a general trend of increasing inequality in the distribution of organizations over European clusters (standard deviation of organizations per cluster more than doubles from 1999 until 2005).

When we look at the evolution of alliance activity of the organizations in these clusters in table 4.1 and Figure 4.3 it is no surprise that alongside the increasing number of deal-active organizations, the number of alliances has increased as well. In figure three, the size of the nodes represent the intensity of local collaboration (within clusters) and the ties between nodes reflect the intensity of non-local collaboration. Two observations are worth mentioning here. One is that while Milan was in 1999

clearly the European centre of collaboration activities, London has gained dominance in 2002 and even reinforced its position in 2005. While this shift from Milan to London is apparent in the alliance network, this is not the case in the geographical landscape of figure 4.2. This indicates that the regional clusters dominating the European landscape in terms of size (representing the number of firms) does not necessarily imply that these clusters are also central in terms of connectivity. A second observation is that it seems at all moments in time, that local link formation and non- local link formation are complements and not substitutes (the largest nodes in figure 4.3 are also the most connected nodes). This observation might hint at the potential of combining visions of clusters as regional growth engines through local collaboration (Storper, 1995; Scott, 1993) with views of clusters as competitive hubs in global networks through international collaboration (Porter, 1990; Batheld et al., 2002). Although our study focuses on explaining local link formation, this preliminary data view informs us of the importance to regard local link formation in relation to link formation in general or to 'international' link formation more specifically.

Fig. 4.2 – European pharmaceutical clusters in 1999, 2002 and 2005. (Node size represents the number of deal-active firms)
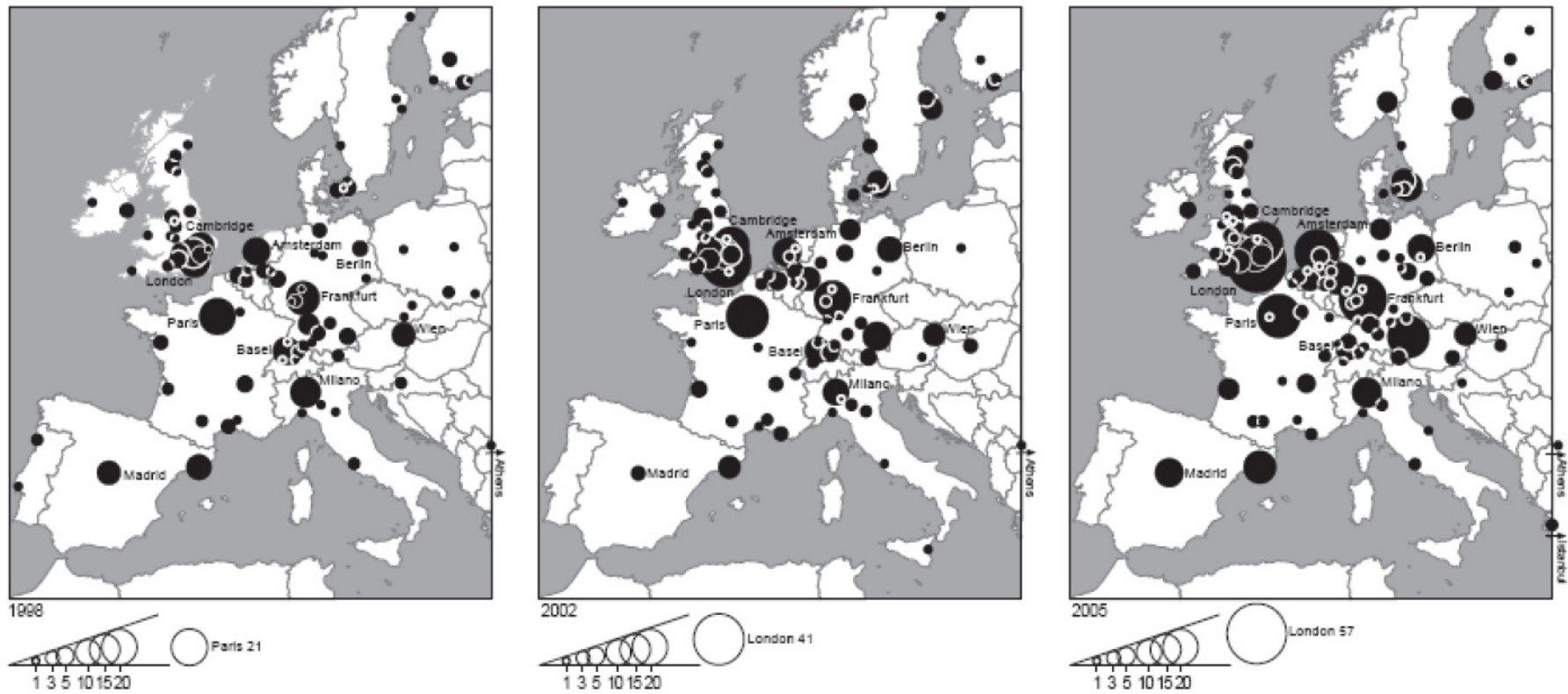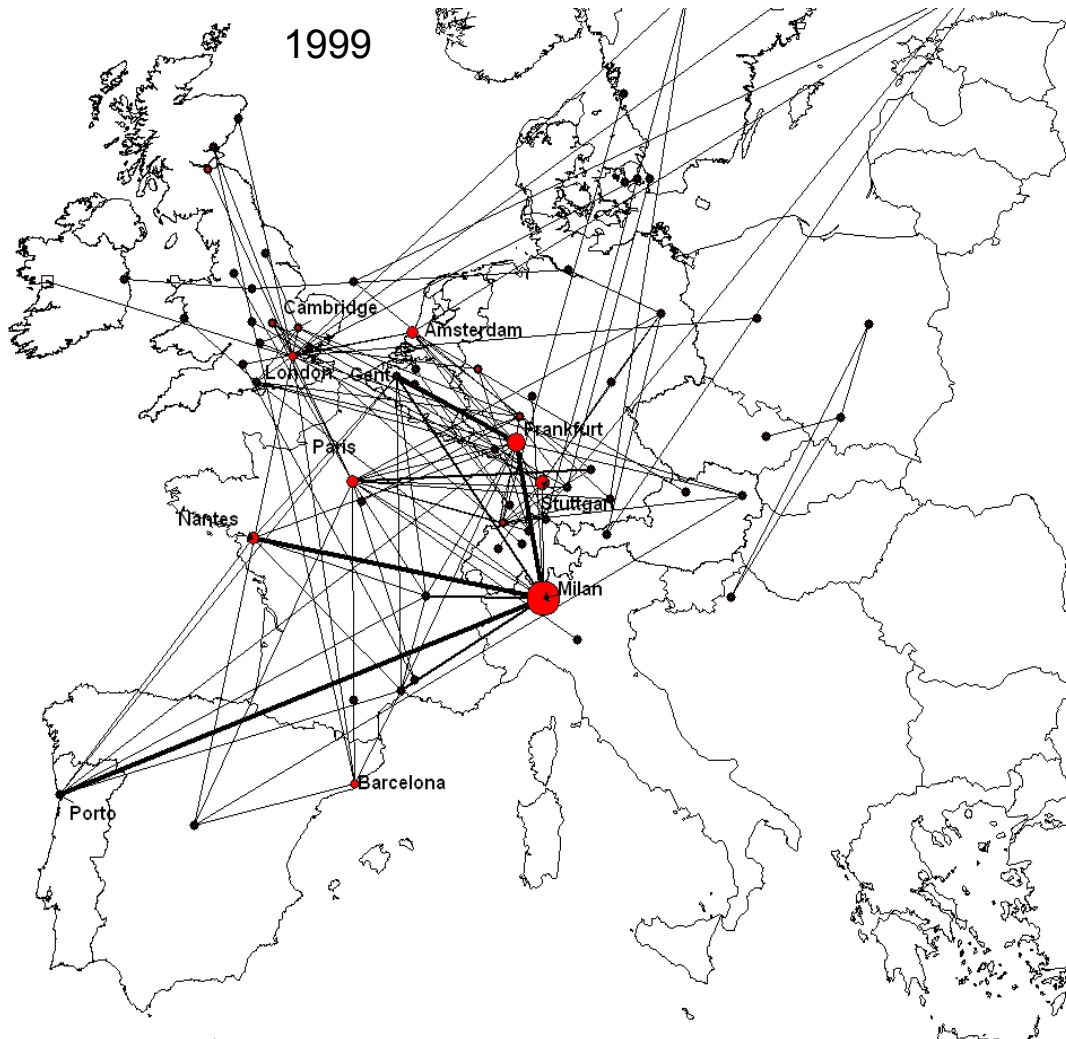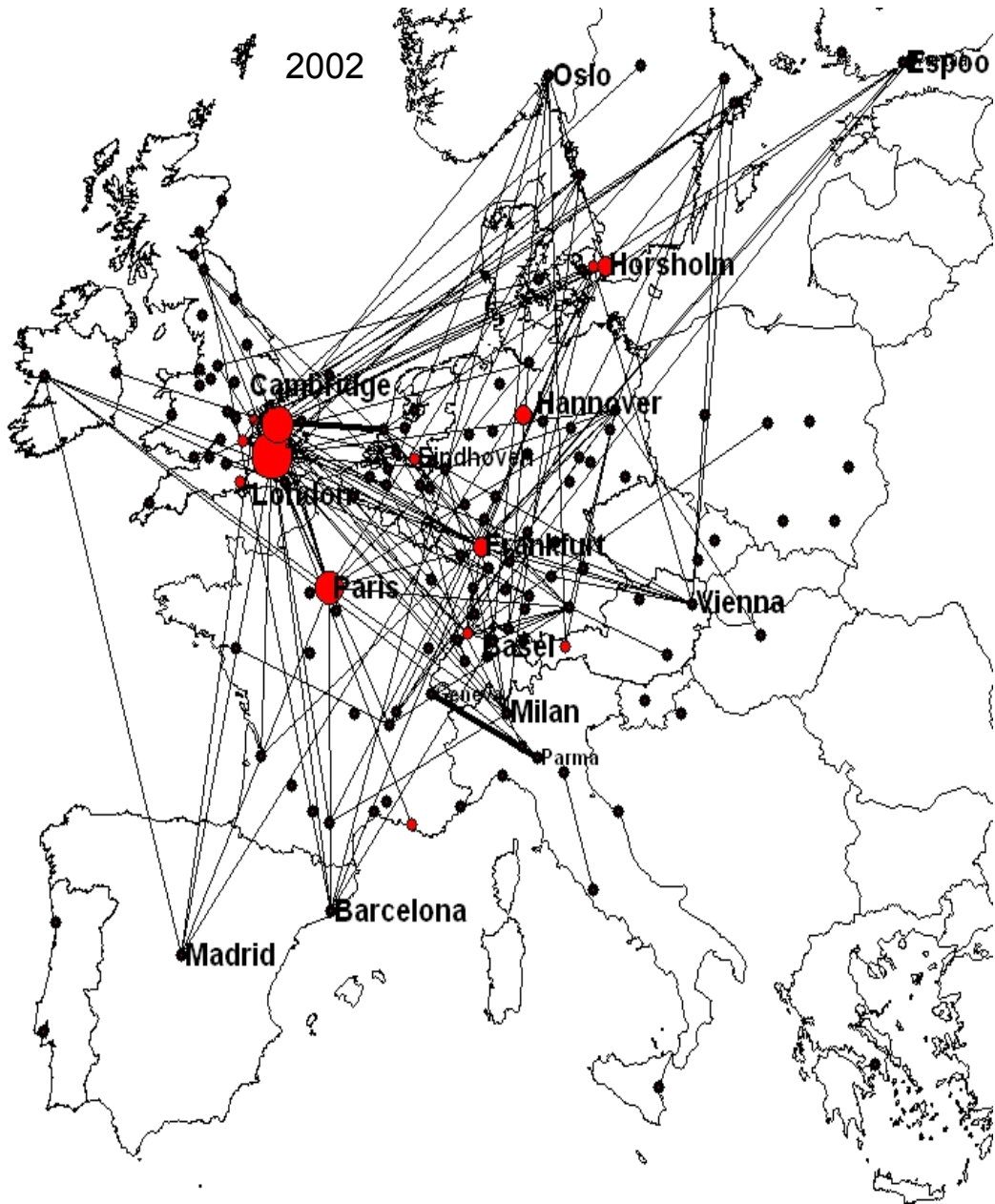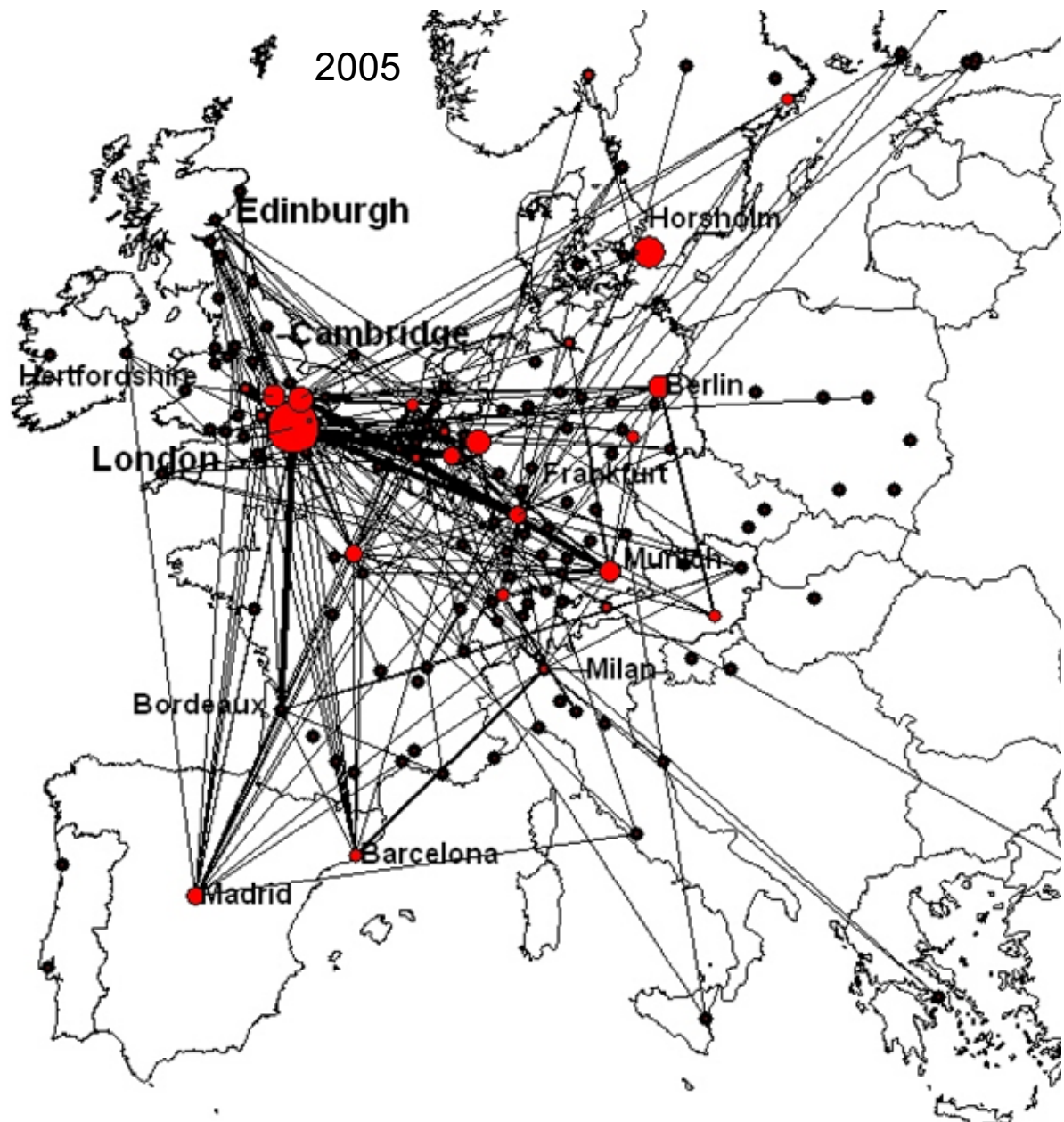
Fig. 4.3 – Collaboration between clusters (ties) and within clusters (node size) in 1999, 2002, and 2005

1999

2002
Espoo
Oslo
Horsholm
Cambridge
Hannover
Eindhoven
London
Frankfurt
Paris
Vienna
Basel
Geneva
Milan
Parma
Barcelona
Madrid

2005

4.3 Literature & Hypotheses

The literature on local knowledge transfer through strategic collaborations is divided into studies emanating from economic geography and management studies focusing on strategic technology partnering. Within economic geography, local knowledge transfer is mainly analyzed as one reason why firms co-locate. Strategic collaboration has in this view been analyzed as an important route through which knowledge is transferred locally. From the perspective of management studies the emphasis is on strategic collaborations and partner choice, whereby geographical clusters are merely considered as a potential place to meet. Because interesting empirical contributions to our research question originate from both disciplines, we aim to integrate the findings under the heading of link formation and local link formation more specifically.

4.3.1 Link formation

The main argument for the formation of inter-firm linkages has been that it provides firms in innovation-driven industries access to new knowledge (Mowery, 1996). Especially in the pharmaceutical industry there is an abundance of empirical evidence of the increasing importance of inter-firm alliances as a way to acquire knowledge that is crucial for innovation (Hagedoorn, 2002; Owen-Smith et al, 2002).

Other reasons for link formation that have been advocated are efficiency reasons (risk & cost sharing, mutual specialization and consolidation of production capacity) and strategic reasons such as the improvement of the firm's long-term product market position. A relatively new argument in management studies originating from sociology is the embeddedness perspective on the understanding of inter-firm link formation. From the relational embeddedness perspective collaborations between organizations are more likely to occur when they are embedded in previous relations. Relational embeddedness creates trust and transparency among organizations, which in turn increases learning performances and reduces costs of mitigating opportunistic behavior (Gulati, 1998; Wuyts, 2003).

With these theoretical perspectives in mind, we now turn our attention to the influence of geography on the process of link formation. In other words, how does 'being in the neighborhood' affect the above described reasons for strategic collaboration? Furthermore, related to the question of why organizations collaborate is the question

of when to expect local collaboration. To answer these questions we build on insights from economic geography, where geographical proximity between organizations and the effect on knowledge flows has been studied intensively.

4.3.2 Local link formation

The nature of knowledge

The competence based view of the firm argues that geographical proximity is beneficial for firms in the process of mutual knowledge acquisition. The reason for this is that geographical proximity offers the possibility of face-to-face contact, which in turn facilitates the transmission of highly specified knowledge (Lundvall, 1992; Saxenian, 1994). Highly specified knowledge (as opposed to codified knowledge) is often embedded in routines, peoples and in machines, and therefore its transfer is geographically bounded (Nonaka, 1994; Blanc & Sierra, 1999; Jaffe & Trajtenberg 1993; Zucker, 1996). In sum, the competence based view proposes that access to specific knowledge can be regarded as a reason for organizations to form local collaborations. From a transaction cost perspective of the firm, there exists however a danger in inter-organizational knowledge flows, that is, unintended knowledge flows easier within geographical proximity and can result in opportunistic behavior (Narula & Santangelo, 2005). From this point of view, knowledge that diffuses as a non-excludible public good (e.g. urbanization economies) can explain the geographical clustering of firms since it reduces transaction costs, but it does not explain local link formation. Taking these arguments together we expect local link formation to occur only when knowledge is excludable (through patents or naturally) and commercially valuable. Furthermore, the expectation of local link formation should according to Rallet and Torre (2000) be refined, since the requirement of geographical proximity might be temporary and does not always necessitate co-location. These authors argue that as the nature of the collaboration between organizations moves from exploration of potential new products or technologies to the exploitation of the product or technology, tasks of organizations involved become more routinized and the need for geographical proximity decreases. The notion that the proximity requirement depends on the nature of activities and phase of the product life cycle is in line with Saxenian's finding in Silicon Valley, where mainly small and medium sized firms explore niche

markets and are at the forefront of new knowledge, which is still non- routinized. Thus, one can say that as far as explorative and exploitative activities in new product development are undertaken by separate organizations, the organizations involved in exploration are expected to be co-located in the same cluster.

In the pharmaceutical industry, the production process of new drug development is roughly divided into explorative or upstream activities aimed at the discovery of new drugs and into exploitative or downstream activities aimed at the development, marketing & distribution of drugs (Powel, 1996; Liebeskind et al, 1996). Existing empirical studies have mainly focused on the organization of upstream activities, since here, naturally excludable and highly specified knowledge is transferred between organizations through formal alliances. Furthermore, it has been shown that in the US at least, these collaborations are formed between co-located organizations (first academic and start-ups firms, later established firms), indicating the importance of geographical proximity for knowledge transfer (Zucker et al, 2002). We now can formulate our first hypothesis as follows:

Hypothesis 1. Clusters with relatively many upstream alliance activities (focusing on exploration) have a higher probability of local collaboration.

Technological proximity

While the above posed hypothesis is founded in an abundance of existing empirical evidence, there is a recent and growing literature in management- and innovation studies that seems to argue otherwise. In these latter studies, partner selection (or link formation) is the main dependent variable that one seeks to explain and geographical proximity is placed next to technological proximity and relational proximity as potential explanatory variables for the formation of a formal alliance (Gulati, 1995; Sorenson & Stuart, 2002). It is argued that in order for firms to effectively collaborate, learn and apply the resulting knowledge in the organization, it is necessary to have a certain minimum level of technological proximity (Wuyts, 2003). Technological (also referred to as cognitive) proximity indicates a common 'mental' framework that cannot be obtained simply through intensive interaction, but requires training and education in the same specific knowledge domains. Given a minimum amount of technological proximity, it is the complementarities or variety in the knowledge bases

of both firms that induces the highest and most effective knowledge spillovers. Organizations that are after these knowledge spillovers will collaborate with partners in any geographical location, and they will more likely use their network of social relations or some professional occasion (such as a conference) to connect to these potential partners. While collaboration between organizations might sometimes be within a regional cluster, it is the technological proximity that drives the relation rather than geographical proximity (Breschi & Lissoni, 2001). Regional innovative clusters such as Silicon Valley or the Boston cluster (around MIT & Harvard) are examples of this latter phenomenon.

When we translate the notion of technological proximity between firms back to the level of a regional cluster, we find that it coincides with both the notion of 'flexible specialization' (Piore & Sabel, 1984) in a cluster *and* with the notion of 'related variety' (Frenken, 2007). 'Specialization' and 'variety' (diversity) represent two opposing views in the economic geography literature as the answer to what induces regional knowledge spillovers. In the literature these are referred to as MAR externalities versus Jacobs's externalities respectively (see Autant-Bernard (2006) for a discussion). By conceptually distinguishing diversity as related-, and unrelated variety (Frenken, 2007) these two (seemingly opposing) views come together. 'Specialization' refers to the relatedness or the similarity of technologies and 'flexible' refers to the benefits that accrue to co-located firms from the complementarities or variety in related activities. Both concepts 'flexible specialization' and 'related variety' basically argue that knowledge spills over between organizations, which are involved in complementary activities with minimum technological proximity. The main difference between both concepts lies in the level of sectoral aggregation at which knowledge spillovers are assumed to occur. While the 'flexible specialization' thesis focuses on complementarities within an industry, 'related variety' is focused on inter-sectoral complementarities. As we will come to see in the data section of this chapter, we propose to let go of any assumptions about the appropriate aggregation level at which spillovers occur and simply test whether complementarities pay-off in terms of local collaboration.

From the above literature discussion we can conclude that being co-located does not in itself provide incentives for collaboration. However, if firms in a cluster are active

in a variety of related activities, we expect to see an increased probability to collaborate. Hence our second hypothesis:

Hypothesis 2. Clusters with a high variety in related technological / knowledge domains have a higher probability of local collaboration.

The type of organizations present

The theory of neo-Marshallian districts states that regional collaborations take place mainly between small firms that are specialized in different parts of the production process of similar products. The smallness of the firms in these districts is perceived essential in order to provide mutual interdependence and trust (Markusen, 1996; Simmie & Sennet, 1999). While not related to a neo-Marshallian district, Saxenian (1991) also finds a more practical reason why small and medium sized firms are more likely to collaborate locally. Start-up firms often collaborate locally in science driven clusters because they are university spin-offs, whose founders stay close to university because of dual occupations. This leads us to a third hypothesis regarding when to expect local collaboration

Hypothesis 3. Clusters with relatively many start-up firms have a higher probability of local collaboration.

The work of Zucker et al (1996) has shown that biotechnology star-scientists at universities are rather entrepreneurial and are responsible for local alliances between the university and local private firms. While for US data there is convincing evidence of cluster formation around academic organizations, there is still very little insight in whether European academic organizations have the same effect. Our fourth hypothesis can therefore be stated as follows:

Hypothesis 4. Clusters where relatively many academic institutions are located have a higher probability of local collaboration.

(Cluster) Life cycle effects

The above literature suggests that local collaboration is more likely to occur when academic and start-up organizations are co-located. However, it might be the case that when too many collaborating organizations are co-located, the information that is exchanged becomes redundant. This, in turn, may reduce the incentives for firms to form new collaborations. Evidence of such decreasing returns is provided by Fingleton et al (2005) in their study on the effect of geographical clustering on high technology SMEs in the computing industry. While sector specific, the commonality of characteristics between computing and biopharmaceuticals, such as being technology and innovation intensive, with rapid growth and the crucial role of SMEs (Fingleton et al, 2005) has been confirmed in the study on cluster evolution by Swann & Prevezer (1996). While redundancy is certainly a plausible explanation for decreasing returns to local collaboration, we will also investigate a second explanation for decreasing returns. It might be the case that as clusters increase in terms of number of organizations, their function changes from providing organizations local 'support' collaborations to being a hub from where organizations connect internationally. In this latter situation, decreasing returns to local link formation should be accompanied by increasing returns to non-local link formation. Or, to put it differently, clusters that grow large enough become (reaching a certain threshold) visible in the global network, which helps organizations to go 'international'. We will test whether a growing number of deal-active organizations in a cluster causes decreasing returns to the propensity to collaborate locally. Additionally, we will investigate whether a decreasing tendency to collaborate locally is complemented by an increasing probability to collaborate non-locally.

Hypothesis 5. There is an inversed u-shaped relation between the number of deal-active organizations and the probability to collaborate locally, indicating decreasing returns to local collaboration.

4.4 Data and methods

4.4.1 Data

We used the Pharmadeals database, which globally monitors the announcement of a variety of strategic alliances in the biopharmaceutical industry on a daily basis since 1996. The data preparation has been a three stage process.

To start with, we selected alliances where at least one European partner was involved. The database did not at that time allow us to create a query based on the location of organizations, and we used European countries as keywords to search for a match in any part of the alliance announcement (including press release). This resulted in a hit of 2800 alliances between June 1996 and June 2005.

In a second stage we have split alliances where multiple partners were involved. Of these 2800 alliances, some 300 appeared to involve more than 2 separate organizations. We decided to split those alliances into each possible combination of alliances. Thus, an agreement with n participating organizations was transformed into n x $(n$-1$)/2$ linkages. Hence, we assume that an alliance serves as a conduit of knowledge transfer, where information transfers between all participants in an alliance (whether the alliance is a R&D collaboration between multiple universities or a project of some pharmaceutical companies who received EU funding). Turning all these multiple partner alliances into dyads led to a new alliance set of 4031 alliances among 2500 separate organizations.

Thirdly we manually looked up the cities where organizations are located and found around 650 organizations either not tractable or undisclosed by Pharmadeals or being a USA based firm with a USA based partner (this latter phenomenon can be explained by the splitting of multiple partner projects).

Our final work set consisted of 2566 alliances among 1834 organizations of which 1054 are organizations located in Europe. We geo-coded each organizations' location and used a hierarchal clustering algorithm based on Euclidian distances between locations to define 54 European clusters. Since we are interested in local collaboration we define a cluster as a geographic area where at least 2 deal- active organizations are located. Having identified the European clusters we aggregated the information on organizations and their alliances back to the regional clusters. Table 4.2 summarizes the terminology and definitions of our clusters.

Table 4.2 – Definitions of cluster attributes

Cluster name = the city of a cluster where most companies are located
Cluster size = the number of deal-active organizations located in a cluster
Alliances (linkages) = the total number of alliances formed by organizations located in a cluster 　　–Intra-cluster alliances= the number of alliances where both partners are located in a cluster 　　–Inter-cluster alliances= number of alliances where a partner is located outside a cluster
Company types per cluster = number of start-up firms, established firms, academic organizations or financial organizations located in a cluster
Therapeutical focus of the alliances of organizations located in a cluster (e.g. oncology, inflammation, etc.)
Type of alliances formed by organizations located in a cluster (e.g. R&D collaboration, manufacturing & supply, co-development, licensing)
Phase in the production process of alliances formed by organizations located in a cluster (e.g. discovery, clinical testing I/II/III, marketing & distribution)

Finally we have created three 'snapshots' of our clusters over time. Following an extensive study by Phelps (2003) on the duration of alliances, the most realistic assumption of alliance duration is three years. The assumption of three year alliance duration is however an average. While our data enable us to distinguishing different alliance types, we found no information on average alliance duration per alliance type. Based on three year alliance duration we have created the following cluster snapshots:

- European clusters at t1 = based on the alliances announced between June 1996 and May 1999.

- European clusters at t2 = based on the alliances announced between June 1999 and May 2002.

- European clusters at t3 = based on the alliances announced between June 2002 and May 2005.

In order to increase our observations we pooled the observations of our explanatory variables of t1 and t2 and our observations of the dependent variable of t2 and t3 respectively. We performed a Wald test to make sure there is significant variation between the information provided by the recurring clusters, so as to treat them as independent observations. This increases our number of observations to 108, as shown in the descriptive statistics (table 4.3) below.

Table 4.3 – Descriptive statistics

	Academic	Established	Financial	Downstream collaborations	Upstream collaborations	Local collaboration	non-local collaboration	Nodes	Variety	Startup	USA connectivity
Mean	0.71	3.09	0.05	4.26	5.29	1.05	15.72	5.80	1.95	1.86	5.10
Median	0.00	2.00	0.00	2.00	1.00	0.00	9.50	3.00	2.03	1.00	3.00
Maximum	10.00	19.00	1.00	35.00	67.00	21.00	109.00	41.00	4.32	18.00	34.00
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.00	0.00	0.00	0.00
Std. Dev.	1.35	3.52	0.21	6.65	11.44	2.75	18.14	6.13	1.20	2.96	6.51
Skewness	3.60	2.13	4.32	2.62	3.67	4.56	2.35	2.90	-0.09	3.10	2.23
Kurtosis	22.20	7.89	19.65	10.38	17.20	29.08	10.24	13.83	2.14	14.45	8.94
Observations	108.00	108.00	108.00	108.00	108.00	108.00	108.00	108.00	108.00	108.00	108.00

4.4.2 Dependent variable

As our aim in this chapter is to explain the determinants of collaboration within a geographical cluster, our dependent variable is intra-cluster collaboration (Local_coll_Y). Intra-cluster collaboration is measured as the number of alliances where both partners are located inside the cluster. While this measure captures only formal alliances that are publicly announced, we argue, following Zaheer and George (2004) that firms in technology intensive environments can only transfer valuable knowledge through formal interactions and not through mere physical presence in a cluster or through informal social contacts. We also argue that possible knowledge spillovers or more informal interaction travels along these formal alliances (during lunch meetings, dinners or other occasions with alliance partners). In order to minimize reversed causality we lagged the dependent variable.

4.4.3 Independent variables

Control variables

Local collaboration at t-1 (Intraclusx). While the issue of reversed causality has been taken into consideration through the lagging of the dependent variable, network studies show that there is another danger, which is autocorrelation. Within the network literature there is general agreement that new alliance formation is a path-dependent process, where firms are likely to partner with previous partners or with partners of previous partners (Gulati, 1995; Coleman, 1988; Granovetter, 1985). Given this feature of alliance behavior, lagging a dependent variable doesn't completely solve the issue of causality, because local collaboration at time t can be driven completely by its local collaborations at time t-1, which could in turn drive our explanatory variables. In order to control for this effect we inserted the number of local collaborations at time t-1 (intraclusX) as control variable.

Cluster size (Nodes). We controlled for size effects of clusters by counting the number of deal-active organizations.

Explanatory variables

Start-up. Number of deal-active start-up firms located in a cluster

Financial. Number of deal-active financial firms located in a cluster

Established. Number of deal-active established firms located in a cluster

Academic. Number of deal-active academic organizations located in a cluster

*Upstream alliances (Linksup).*With this variable we aim to capture explorative knowledge transfer (including highly specified knowledge). For each cluster we have counted the number of alliances aimed at drug discovery or R&D. We further included alliances that involved academic organizations, and alliances that were funded by the government (as basic science collaborations often are).

Downstream alliances (linksdown). With this variable we aim to capture exploitative knowledge transfer (including codified knowledge). For each cluster we have counted alliances that where active in marketing, distribution, manufacturing & supply or in co-promotion activities.

(Related) variety. In our second hypothesis we proposed that clusters in which there is a higher variety of related technological- or knowledge domains, we expect a higher probability of local collaboration. Since the information of our clusters is build upon alliance information and not upon organizational information, we have to rely on a proxy for technological proximity. We used the therapeutic focus of an alliance formed by organizations located in a cluster as if it were an organizational characteristic. Our justification for this strategy is that organizations can only collaborate in a therapeutic area where they are knowledgeable. We are however aware of the misinterpretation danger of this measure, since organizations could also collaborate in knowledge domains, which are complementary to them. Support for our approach can be found in the relational view of the firm, which states that a firms' portfolio of alliances can be regarded as an important resource of the firm and a major source of competitive advantage (Dyer & Singh, 1998). Following a number of previous studies that aim to measure technological proximity or related variety in a geographic area, we calculated the degree of entropy of each cluster (Frenken 2007; Boschma & Iammarino 2007). While Boschma & Iammarino (2007) further specify their definition of variety to capture related variety, we can measure related variety

using their variety measure. The reason for this is that our data are based on alliance information, which by definition already ensures relatedness. Furthermore, alliances are aimed at fulfilling a part of the drug development process, which implies some degree of technological or cognitive proximity. Variety within these activities is caused by linkages focusing on different therapeutical areas and even on technologies originating from different sectors such as informatics (e.g. bioinformatics), or mechanical engineering (e.g. medical devices). Formally we measure (related) variety as entropy, which we define as:

$$\text{Variety} = \sum_{i=1}^{N} p_i \log_2\left(\frac{1}{p_i}\right)$$

pi represents the share of linkages formed in therapeutic area *i*, with *i* ranging from 1 to *N* in each cluster. As the range of therapeutic areas in a cluster is partly driven by the size of a cluster, we weighted our clusters by size. Variety thus measures the weighted sum of therapeutical areas in which organizations are deal-active whereby therapeutical areas with relatively fewer linkages are weighted stronger.

Nodes & Squared nodes. Measuring the effect of cluster size (nodes) and the squared cluster size (squared nodes) on local link formation enables us to see whether there are decreasing returns to local collaboration. The interpretation of decreasing returns to link formation can be twofold: First, it might be that the cluster has become so crowded with interacting organizations that the information they transfer has become redundant. Second, the number of organizations in a cluster may increase up to a point where the cluster becomes 'internationally visible' and, as a result, induce firms to collaborate internationally instead of locally. If the latter scenario holds, we expect to see decreasing returns to local link formation in combination with increasing returns to non-local link formation. Variables that were distributed non-normally have been transformed to logarithms.

Non-local linkages. While our aim is to understand what causes local link formation, there are two reasons to take non-local link formation into account. First, if we want

to be sure that factors, which we find to induce local link formation are specifically local or whether these factors induce link formation in general, we need to compare the effect of our predictors on local and on non-local link formation. Second, as we have seen from our preliminary data view, regional clusters may perform a dual role in an economy, namely they may be locally embedded and at the same time function as nodes in a global network. While the investigation of this dual role is beyond the scope of this chapter, it would be naive to confine our view on inter-firm collaborations to local collaborations, especially given the global scope of the pharmaceutical industry.

4.4.4 Analysis

We test our hypotheses using negative binomial regressions. While Poisson models are often used with count data, these models require the mean to be equal to the standard deviation. As the descriptive statistics in table 4.3 show, this latter requirement is not fulfilled in our dependent variable. As our dependent variable contains a high number of zero values (indicating that no local collaboration takes place), we performed a Vuong test to see whether using a zero-inflated negative binomial regression model would better fit our data. This was not the case.

4.5 Results

Table 4.4 shows the regression results, which are presented separately for each hypothesis.

Table 4.4 – Negative binomial regression results using maximum likelihood estimations

	Dependent Variable: local link formation				**Dependent Variable: non-local link formation**			
	Coefficient	Std. Error	Z-Statistic	Prob.	Coefficient	Std. Error	z-Statistic	Prob.
Effect of up- & downstream activity in a cluster on local (& non-local) link formation								
C	-5.280936	0.95	-5.55	0	0.720072	0.15	4.855	0
LOG(LINKSUP+1)	-0.263489	0.23	-1.13	0.26	0.465	0.04	12.81	0
LOG(LINKSDOWN+1)	-0.006921	0.24	-0.03	0.98	0.449	0.06	7.985	0
LOG(NODES+1)	2.630704	0.5	5.215	0	0.371713	0.1	3.812	0.0001
Effect of related variety in a cluster on local (& non-local) link formation								
C	-7.5182	1.99	-3.78	0.00	0.52756	0.24	2.14	0.03
variety	1.025551	0.60	1.70	0.08	-0.174492	0.04	-3.53	0.00
VARIETY*LOG(NODES+1)	-0.461937	0.30	-1.52	0.12				
LOG(NODES+1)	3.548979	1.22	2.89	0.00	1.359283	0.17	7.89	0.00
Effect of organization type in a cluster on local (&non-local) link formation								
C	-3.293152	0.68	-4.88	0	1.348962	0.16	8.614	0
LOG(ACADEMIC+1)	0.624843	0.36	1.747	0.08	0.649114	0.12	5.489	0
LOG(STARTUP+1)	1.108469	0.26	4.258	0	0.348737	0.1	3.634	0.0003
LOG(ESTABLISHED+1)	1.161943	0.31	3.798	0	0.588432	0.09	6.237	0
Decreasing returns to local (&non-local) link formation								
C	-6.796358	1.65	-4.12	0	0.618971	0.35	1.761	0.0782
LOG(NODES)	5.016013	1.54	3.258	0	1.710182	0.51	3.344	0.0008
LOG(NODES)^2	-0.748472	0.38	-1.95	0.05	-0.25505	0.17	-1.54	0.1238
INTRACLUSX	0.078268	0.13	0.587	0.56	0.106859	0.04	2.444	0.0145
USA effect on local (&non-local) link formation								
C	-0.741031	0.34	-2.19	0.03				
CENTERED_USA	-0.016835	0.05	-0.32	0.75				
CENTERED_NODES	0.24071	0.07	3.681	0				
LOG(ACADEMIC+1)	0.027102	0.42	0.065	0.95				
LOG(INTRACLUSX+1)	-0.190417	0.43	-0.45	0.66				

4.5.1 The nature of knowledge

The first set of variables relate to our hypothesis that upstream alliance activity, which entails the exchange of highly uncertain and specified knowledge (such as drug discovery) is positively related to local link formation. To see whether this relation is specific to upstream alliance activity, we compared the effects of upstream alliance activity to the effects of downstream alliance activity. Further we tested whether our hypothesized effects are specific to local link formation by comparing the effects of each of our explanatory variables on local link formation with their effect on non-local link formation. Our findings lead us to reject our 1st hypothesis indicating that upstream, explorative knowledge exchange does not have a significant effect on local link formation. However, we do find that upstream knowledge exchange strongly and positively ($z = 12.81$) predicts expected non-local link formation. This result seems to indicate that valuable and highly specified knowledge is sourced from any location. However, we have to be very careful since not only upstream alliance activity but also downstream alliance activity is positively correlated with non-local link formation.

4.5.2 Related variety

Our second hypothesis stated that clusters with a higher variety in related technologies (therapeutic areas) have a higher probability of local link formation. The more variety in a cluster, the higher the potential knowledge spillovers and learning effects would be. While we found no significant effect of variety on local link formation for the whole sample, we found a positive and significant effect of variety for relatively small clusters. We distinguished between variety effects in small clusters and large clusters by adding an interaction variable between variety and cluster size. The positive effect of variety in smaller clusters is in line with the intuition that small clusters provide organizations with local support while large clusters function as global hubs. Local support in these clusters could be provided through inter-organizational collaboration if these organizations' activities are of a complementary nature (indicating a high variety). Further support for the above described intuition arises when we take a look at the effect of variety on non-local link formation. Variety has a negative significant effect on non-local link formation. This means that in clusters where activities are more specialized (less variety), there is more 'international' (non-local) link formation. Again, the intuition seems in place that clusters, which function as global

hubs search for complementary knowledge through 'international' alliances and do not benefit from a local variety of knowledge or technologies. In sum, we can say that in small clusters firms form more local 'support' collaborations when there is more therapeutical variety. When firms in a cluster are active in similar therapeutic areas (specialization) they are capable of forming more non-local or international collaborations.

4.5.3 Type of organizations present

Hypotheses 3 and 4 relate to the type of organizations present and whether or not to expect local link formation from their presence. Table 4.4 shows the estimated effects of academic organizations, start-ups firms and established pharmaceutical firms on local link formation. Financial organizations are treated as a reference category. Academic organizations are the least strong drivers of local collaborations, while they are the strongest drivers of non-local ('international') linkages. This finding leads us to reject hypothesis 3, which stated that academic organizations collaborate locally. This result is in line with our previous finding that explorative activities, which are often carried out by academic organizations are non- local. Our fourth hypothesis, regarding the positive effect of startup firms can be confirmed. Startups are more likely to collaborate locally and local collaborations are more likely to occur than non-local collaborations. Finally, while we did not hypothesize on the effect of established firms based on previous literature, established firms do perform an interesting role in link formation. While academic organizations mainly induce non-local collaborations and startups are more likely to collaborate locally, established firms are the drivers of *both* local *and* non-local linkages. In other words, established firms can be said to exert a bridge function in connecting a cluster locally and 'internationally'.

4.5.4 Cluster life cycle

Our last hypothesis (5) deals with the issue of the life cycle of a cluster. The assumption in the literature is that as the number of active organizations in a cluster grows, there will be decreasing returns to local link formation. We proposed two testable views to explain this phenomenon. One is that congestion effects occur because the information exchanged through collaborations is becoming more and more redundant as the cluster grows. A second scenario we investigated is one where

larger clusters are more globally 'visible', enabling firms to collaborate non-locally rather than locally. If this latter scenario is indeed true, we expect to find increasing returns to non-local link formation. Based on our findings in table 4.4 we confirm decreasing returns to local link formation, but reject increasing returns to non-local link formation[3]. We interpret this finding as evidence of local congestion effects. In other words, as the number of organizations in a cluster grows, the number of new collaborations, both local and non-local, decreases.

4.6 Discussion & conclusion

The primary aim of this chapter was to help policy makers understand why some regional clusters are networks of intense collaboration while most clusters are mere co-located firms. Behind this aim lies an extensive literature that claims these networks of intense collaboration are a crucial factor for the innovative performance of firms that make up these clusters. Our findings suggest that indeed only very few regional clusters can be described as networks of intense collaboration. However, this skew distribution of linkages over clusters is not specific to local link formation but characterizes alliance activity in general. In other words, there are some clusters that are networks of internal collaboration, which are at the same time important hubs in non-local networks while most clusters are merely co-located firms. Moreover, most other cluster characteristics that we investigated are also skewly distributed, which indicates the high diversity of European pharmaceutical clusters. This might explain the prevalence of case studies in this area focusing on one or two specific clusters.

There are two main lessons that policy makers can take from our analysis. The first lesson is a rejection of the most common ideas about when to expect local link formation, and the second contains some clear evidence of factors that do induce knowledge exchange through local collaboration.

First, it is rather striking and counterintuitive that whereas in the US highly specified explorative knowledge spills over from universities into US clusters through formal collaborations, in Europe this appears not to be the case. Instead, explorative R&D activities such as early stage drug discovery induce non-local alliance formation. This

---

[3] Decreasing returns to local link formation is indicated through a switching coefficient (positive to negative) for *nodes* and *squared nodes* respectively. A simple scatter plot (appendix I) shows that the probability to collaborate locally does not become negative within our range of organizations.

finding is in line with what Adams and Jaffe already found in 1996, namely that pharmaceutical R&D spillovers do not decrease as much with distance as other industries' R&D does. One reason for the non-stickiness of pharmaceutical R&D in Europe could be that research centers and start-ups with 'forefront' knowledge are located in the US, thus making global link formation more important than local link formation for acquiring access to valuable new knowledge. To see whether alliance activity with US partners functions as a substitute for local knowledge transfer, we briefly tested the effect of 'USA connectivity' (number of alliances of firms in a cluster with US partners) on local link formation (see table 4.4). As the table shows, there is no significant effect of USA connectivity on local link formation, and hence we can not speak of a substitution effect between knowledge sourcing from the US and local knowledge spillovers (through alliances) in Europe.

Second, we found clear positive effects of (related) variety in small clusters, of start-up firms and of established pharmaceuticals on local knowledge exchange. These findings support a resource based view on strategic alliances, where local alliances are formed by firms to obtain access to complementary knowledge. While clusters are small and contain relatively many start-ups firms, resource interdependence outweighs the danger of unintended knowledge flows. The danger of unintended knowledge flows, which might increase the transaction costs of collaborating, is further mitigated through the technological variety amongst the collaborating organizations. The positive effect of variety is however specific for local collaboration in relatively small clusters. As we have seen in figure 4.3, the pharmaceutical industry is a truly global industry in which explorative new knowledge is mainly sourced non-locally. For policy makers, this implies that clusters can be growth engines through local collaborations, but need to be present as hubs in non-local networks as well. Our findings suggest that large established pharmaceutical firms are important for achieving this dual goal since they induce both local and non-local collaborations.

Appendix I

Scatter plot of cluster size (nodes) and local collaboration (local coll_y)
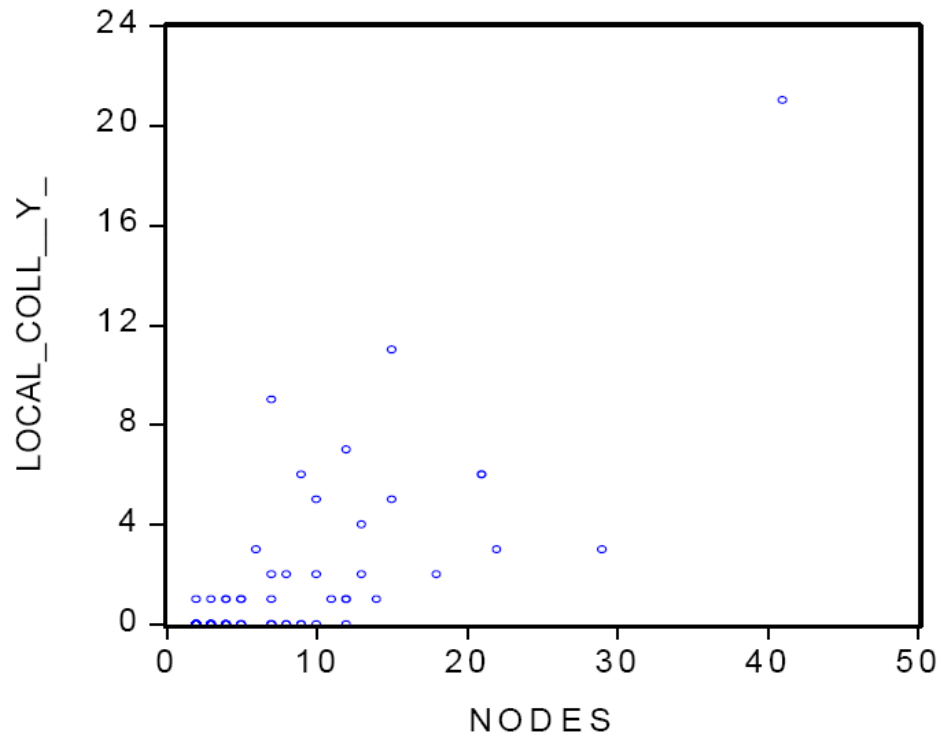
Chapter 5   Radical innovation and network evolution[1]: the effect of the genomic revolution on the evolution of the pharmaceutical R&D network[2]

5.1 Introduction

Networks of collaborative relationships among firms are an important form of organization of innovative activities (Powell et al., 1996; Kogut, 2000). Especially in innovative-, and technology intensive industries, firms increasingly realize that, in order to tap into new technologies and know- how, internal development needs to be complemented with strategic collaborations (Gulati, 1998; Verspagen & Duysters, 2004). In the biopharmaceutical industry, the emergence of an expanding network of R&D collaborations has been studied intensively (Pisano, 1991; Arora and Gambardella, 1994; Powell et al., 1996; Stuart et al. 1999; Orsenigo et al., 2001; Pammolli et al. 2001; Riccaboni and Pammolli 2002).

In this chapter we study the network of innovators in the biopharmaceutical industry to provide a deeper understanding of the underlying mechanisms that drive network structures. We argue that real world networks do not evolve in isolation, but co-evolve along technological paradigms. Our data cover a time frame of 20 years in the biopharmaceutical industry in which a new technological paradigm is established over an old one. The technological transition has been induced by a radical scientific innovation. This has given us the opportunity to explore the effect of radical technological change on structural network evolution.

While radical technological change occurs regularly and has an important influence on the structure of high technology industries and networks (Anderson & Tushman, 1986), we have found that most inter-firm network studies assume the underlying technology base to be stable. Through this assumption, the majority of these studies assume network evolution to be an endogenous process where network structures

---

[1] This chapter is based on joint work with Massimo Riccaboni.
[2] Forthcoming in *Annals of Economics and Statistics*.

guide organizational action and vice versa. In this chapter we provide additional insights into network evolution, because exogenous influences can lead to relational behavior that cannot be explained from an endogenous perspective on network evolution. Another, more theoretical contribution of this chapter relates to the notion of small worlds in complex networks. While most recent studies on small-world networks reveal the existence and topology of such a network structure, this chapter is, as far as we know, the first study to look at influences that might induce the relational behavior that leads to a small-world structure.

We find that the genomic revolution, representing an exogenous shock, leads to an expansion of the network through a wave of firm entry and a wave of alliance formation. When looking more specifically at the partners of new alliances, we find that it requires an exogenous shock like the genomic revolution, for firms to leave their embedded path of existing collaborations and ally with new partners. For managers of new firms or peripheral firms, knowledge about such network changes creates an opportunity to potentially improve one's position in the network. Especially in industries characterized by hierarchical structures and low turnover, where large incumbent firms are dominant and newcomers are usually specialized niche players, knowledge about such structural breakthroughs can be crucial. While it is interesting for firms to be able to anticipate the effect of radical innovation on their and their competitors' position in a network, policy makers can also benefit from this knowledge. Since radical scientific innovation is often induced by government-led R&D programs, policy makers should know about the potential effect of publicly financed R&D projects on firm behavior and network development.

In what follows we first elaborate rather extensively on the existing literature, because our aim is to connect insights from organization-, and strategic management literature to issues in complex network theory. At the intersection of these two strands of literature we find relational behavior that can only be explained when taking exogenous influences into account. In section 2, we introduce our measurement techniques for measuring structural network change, such as the clustering coefficient and the average path-length. Section 3 focuses on exogenous influences on the network. In this section we introduce the reader to our empirical setting, the genomic revolution. We apply and extend a theoretical framework developed by Koka et al. (2006), which enables us to hypothesize on the effect that the genomic revolution has

had on the pharmaceutical R&D network in section 4. We divide this section into hypothesized structural change at the level of the network, and into structural change at the level of the firm, whereby this latter part touches upon specific relational behavior that is associated with small worlds. After describing the data and methodology in section 5, we present the results in section 6 and a discussion and conclusion in section 7 and 8 respectively.

5.2 Network evolution

The understanding of how networks evolve has been a topic of interest to both social scientists and natural scientists. Both scholars realize that relational behavior and network structure are intertwined. An actors' relational strategy depends for some part on the structure of relationships it had before. At the same time, the actors' new relationships contribute to a changed network structure that again influences its actions (Gulati, 1998). Social scientists and natural scientists differ in regard to the way they study network evolution. Social scientists assume that actors conduct strategic relational behavior, while natural scientists, studying complex network theory, often assume actors to be non-human (e.g. proteins). Amongst the latter, the process of link formation is based on certain 'rules of attachment'. Recently, a number of authors such as Goyal et al. (2006), Uzi & Spiro (2005), Wilwhite (2001), Verspagen & Duysters (2004), and Jackson (2006) have combined insights from both strands of literature and increased our understanding of network evolution. This chapter aims to contribute to this understanding.

Network change consists of changes in the number of actors (exit and entry), and changes in numbers-, and patterns of link formation (Koka, 2006). Structural network change is a form of network change whereby new linkages are formed with new partners. Studies on new partner search in networks have broadly focused on two issues. One issue is about distribution of linkages among actors in a network, which represents the inequality of access that firms have to various resources. In many real world networks the distribution of linkages among actors is highly unequal (Dorogovtsev and Mendez, 2003; Goyal, et al.2005; Barabasi et al., 2002). Barabasi

(1998) shows how actors accumulate new linkages in proportion to the number of linkages they have already (preferential attachment). Following from this 'rich get richer' principle of growth, the resulting network structure consists of a few highly connected actors called 'stars' in combination and many weakly connected 'peripheral' players. The second issue in new partner search concerns the process of local link formation and the process of distant link formation, which will be the focus of this study.

5.2.1 Local link formation

Local link formation implies that new partners are found through an actors' existing network (which is called an ego network), and that the new partner is already known to other partners 'in the neighborhood'. The overall network structure resulting from local link formation is a network composed of dense 'cliques' of actors, which indicates that they are highly connected to each other. Local link formation of an actor and the degree of clique formation in a network can be measured by calculating the clustering coefficient, which will be explained in section 2.

Within the organization-, and strategic management literature, network studies mainly focus on the effect that a given network structure has on the relational behavior and performance of firms (Burt, 1992; Granovetter, 1985; Gulati et al., 2000). Having a more central and autonomous structural position in a network provides firms with access to resources, learning opportunities, and reduces uncertainty (A notion first coined by sociologists like Bourdieu (1980) and Coleman (1990) as 'social capital'). Regarding the formation of new linkages, Gulati (1995) finds that the process of new tie creation is heavily embedded in an actors existing network (consisting of previous alliances). This means that new ties are often formed with prior partners or with partners of prior partners (Gulati, 1995), indicating network growth to be a 'local' process, where strategic collaborations are path-dependant (Noria, 1992). Particularly when considering inter-firm alliances, new link formation is considered 'risky business' and actors prefer alliances that are embedded in a dense clique were norms are more likely to be enforceable and opportunistic behavior to be punished (Gulati 1995; Powell et al., 1996; Koka et al., 2006; Granovetter, 1985).

5.2.2 Distant link formation

Distant link formation implies that new linkages are created with partners whom are not known to the existing partners of an actor. In the social sciences, Granovetter (1985) was the first to differentiate between local ties in dense cliques (strong ties) and distant ties that bridge these cliques (weak ties). More precisely, the author argues that distant linkages "serve as crucial functions in linking otherwise unconnected segments of the network" (Granovetter, 1983: 217). At the level of the firm, Burt (1992) shows that distant linkages that serve as bridges between dense local cliques of firms, can provide access to new sources of information and favorable strategic negotiating positions (termed 'structural holes'), which improves the firms' position in the network and industry.

5.2.3 Small worlds

The first network studies that combine local- and distant link formation originate from complex network studies. Watts and Strogatz (1998) model the process of local link formation and find that, with the addition of just a handful of distant linkages, a specific network structure is generated, which they call a small world. This means that although large networks have relatively few linkages compared to the number of actors, the reach is higher than expected (Newman, 2001). While solely local link formation results in dense cliques of connected actors, the average distance to reach all actors in a network is very large. The distance between two actors is indicated by the number of other actors one has to surpass in order to reach the other. Watts & Strogatz (1998) found that the average distance between all actors in a network is sharply reduced when a relatively small number of distant linkages (referred to as random linkages) are added to the network that serve as shortcuts between these local cliques. Examples of small world networks are the electronic power grid network, high-school friendship networks, or the neural network of a worm (see for an overview Watts (1999) or Newman (2001)). Recently, insights from the social sciences regarding network evolution and new link formation have been combined with the more theoretical findings as described above (see Goyal, et al., 2006; Uzi & Spiro, 2005; Wilwhite, 2001; Verspagen & Duysters, 2004). Verspagen & Duysters (2004) explain how firms that try to build 'social capital' can be seen as drivers of

local link formation, and firms that strategically aim to bridge structural holes in a network can be seen as drivers of distant link formation. Together, these two drivers of new partner search add up to small world structures in networks of technology alliances. A recent study by Jackson & Rogers (2006) focuses on link formation with new partners in social networks. They find that large social networks evolve into small worlds, because people meet friends of friends and strangers. The process of link formation is generated by an algorithm that makes actors form both local linkages and random linkages (distant linkages), while implying that random link formation resembles the 'meeting of strangers'. While the latter implication seems feasible in friendship networks, random partner search by firms seems rather unlikely. A plausible assumption made by Verspagen & Duysters (2004), is that firms are aware of the structural features of the network surrounding them and that this induces these firms to deliberately form distant linkages that bridge local cliques. However, a number of studies such as the work of Cowan et al. (2004) and Powell (1990) emphasize the risks involved in new partner formation, and the strong tendency of these firms to use their existing network as a source of information for new partner search, implying local link formation.
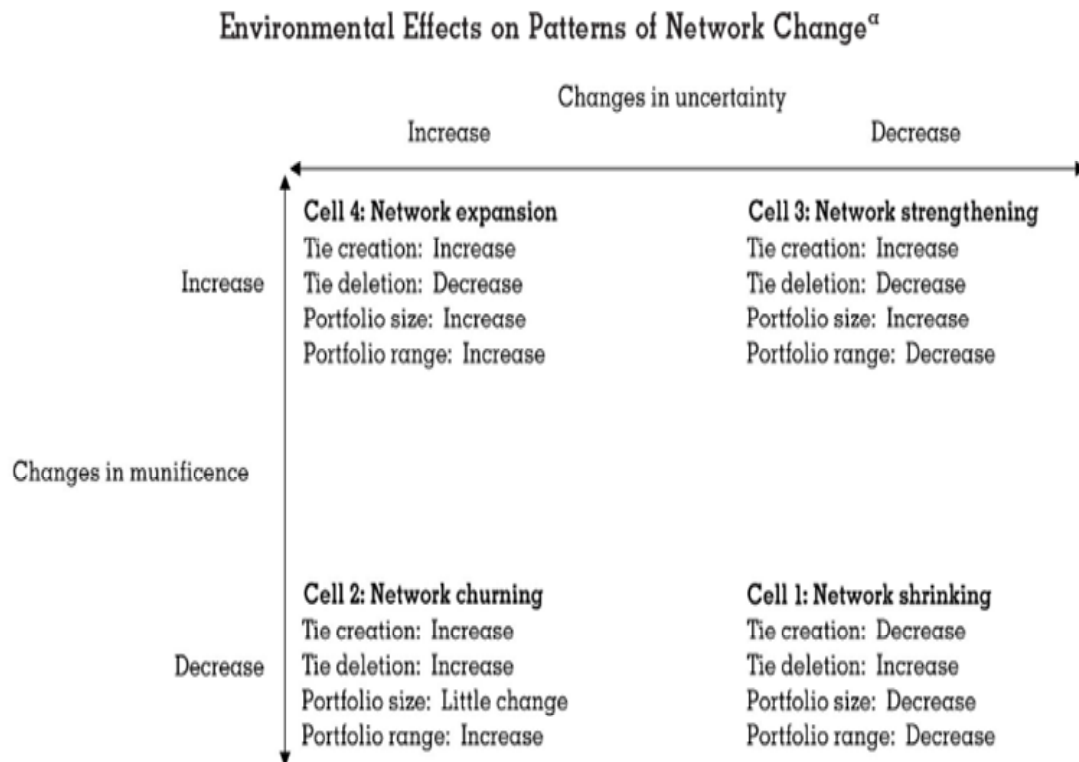
In this chapter we aim to contribute to the understanding of local-, and distant link formation in inter-firm alliance networks. Moreover, we investigate the effect of a radical exogenous innovation on structural network change. This chapter differs from the studies mentioned above, in that these studies analyze the effect of link formation on the emergence of a small-world. Our study investigates whether radical exogenous change induces link formation, which potentially leads to a small world. We use the theoretical framework of Koka et al. (2006) to measure structural change, and we will expand their framework by introducing local-, and distant link formation.

5.3 Structural change

Koka et al. (2006) have combined multiple indicators of relational behavior into four different types of network change (see figure 5.1). The network can expand, churn, strengthen or shrink. Each network change is brought about by a specific combination

of changes in tie creation, tie deletion, and by changes in an actor's portfolio size (number of links) and portfolio range (number of partners).

Fig. 5.1 – Environmental effects on patterns of network exchange

Environmental Effects on Patterns of Network Change[α]

		Changes in uncertainty	
		Increase	Decrease
Changes in munificence	**Increase**	**Cell 4: Network expansion** Tie creation: Increase Tie deletion: Decrease Portfolio size: Increase Portfolio range: Increase	**Cell 3: Network strengthening** Tie creation: Increase Tie deletion: Decrease Portfolio size: Increase Portfolio range: Decrease
	Decrease	**Cell 2: Network churning** Tie creation: Increase Tie deletion: Increase Portfolio size: Little change Portfolio range: Increase	**Cell 1: Network shrinking** Tie creation: Decrease Tie deletion: Increase Portfolio size: Decrease Portfolio range: Decrease

Source: Koka et. al (2006)

While Koka et al. (2006) present four types of network change they find that only an expanding network and a churning network are a reflection of structural change, because new alliances are formed with new partners. An expanding network is brought about by an increase in new alliances without deletion of old alliances (meaning a larger average portfolio), together with an increasing portfolio range (more different partners). A churning network reflects the formation of new alliances and the deletion of existing alliances. While the average portfolio remains stable in terms of the number of partners, there is an increasing variety in identity of partners. We will use this framework to hypothesize on the 'type' of network change to expect after a given exogenous or environmental change. While changes in the number of linkages (tie creation/deletion) and changes in the number and identity of partners already provides important insights into structural changes in the network, we will

further distinguish between local link formation and distant link formation when studying new link formation with new partners. Local link formation and distant link formation are measured through the calculation of the clustering coefficient and the average distance between actors respectively.

5.3.1 Clustering coefficient

The clustering coefficient of an actor is the density of its open neighborhood, that is to say how close each actor's neighborhood is to a fully connected clique. Following Watts and Strogatz (1998), we define a clustering coefficient as follows: assume that the ith vertex v_i has $k_i - 1$ neighboring vertices. At most, $k_i(k_i - 1)/2$ edges can exist between them. Calculate $c_i \equiv$ (number of edges of v_i and its neighbors) $/ k_i(k_i - 1)/2$. Then, the overall clustering coefficient is defined as:

$$CC = \frac{1}{N} \sum_{i=1}^{N} c_i$$

CC is the average of the individual clustering coefficients c_i. The weighted overall clustering coefficient (WCC) is the weighted mean of the clustering coefficient of all the actors each one weighted by its degree. This last figure is exactly the same as the transitivity index of each transitive triple expressed as a percentage of the triples in which there is a path from i to j[3]. The cluster coefficient tends to 1 if most of the partners of each biopharmaceutical institution are directly related by formal R&D collaborations. On the contrary, the clustering coefficient tends to 0 if the network is hierarchical and the partners of each biopharmaceutical actor are not related.

Clustering coefficients are often applied to detect small-world networks and the degree of hierarchy of local relational structures. At the level of the network, the degree of hierarchy of local relational structures is called transitivity. If an actor's

---

[3] To calculate C it is important to notice that there are no loops attached to a vertex (no self-ties) and that multiple relationships between two vertices are identified as one edge.

ego-network is a clique, meaning the absence of hierarchy, the actor and its partners all have equal structural power and the network as a whole becomes more transitive. Because the clustering coefficient (CC) essentially measures the situation whether an actor's partners are connected to each other, an increasing CC measures new alliances being created with partners that are already known to the partners in the clique. We can visualize this alliance behavior as follows:

Fig. 5.2 – Visualization of local link formation

t = 1   t = 2
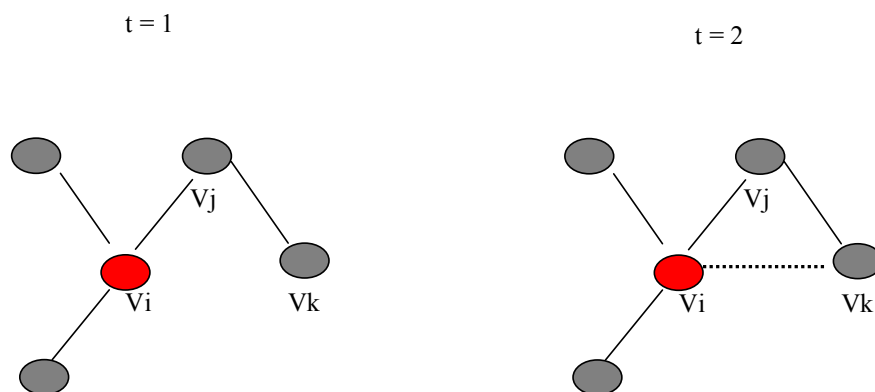
In figure 5.2 we can easily see that whereas at t = 1 vertex *i's* partners are unconnected, the alliance between vertex *i* and vertex *k* at time 2 indicates that *i* has a new partner who already was a partner of *j*. The clustering coefficient has risen from 0 (0/3) to 0.167 (1/6). To account for the size of a firms' ego network, we will use changes in the weighted clustering coefficient as an indicator of structural network change, because it reflects the formation of new alliances with new 'local' partners whom are already known to the actors in a clique.

5.3.2 Average path-length

To measure structural network change that is caused by new link generation *between* local clusters, we use the average path-length in the network. Following Goyal et al. (2006), we define the average path-length between reachable pairs in our network as *d(N)*, being the average distance between any actor *i* and *j* that belong to the same component. Actors that are isolated from other actors are excluded. Thus, for a

connected graph $N(E,V)$ consisting of edges and vertices, the average path-length is the sum of the distances between two actors (i and j) belonging to the network (N), divided by all possible edges excluding self-ties:

$$d(N) = \frac{\sum_{j \in V} \sum_{i \in V} d(i, j; N)}{v(v-1)}$$

The above definition of average path-length is only useful when most actors in a network under study do indeed belong to the same component. To verify this, we plotted the number of actors within the main component of our network in relation to the all actors in the network in figure 5.3. Clearly, this figure shows that at any moment in time, most actors belong to the main component in which every actor is connected.

Fig. 5.3 - Size of the main component (dotted line) in relation to the whole network (full line)
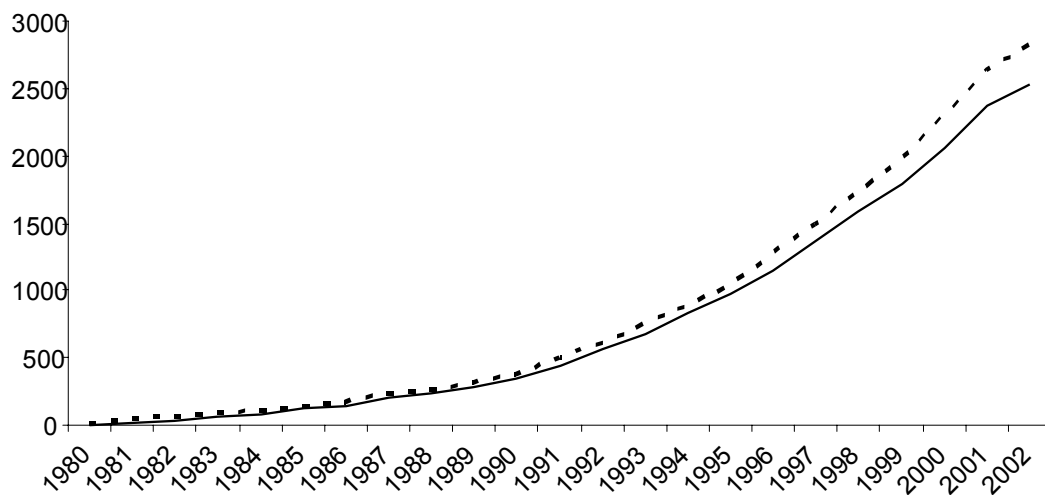
While average path-length is an indicator of the overall network structure, the change in average path-length over time provides information of structural network change at the intermediate level. Moreover, in social networks where the majority of link formation occurs within the neighborhood, only the occurrence of a relatively few number of distant linkages between neighborhoods can cause the average path-length

in a network to fall. This is because a connection (bridge) between to isolated clusters of actors suddenly increases the reach between those clusters and thus decreases the path-length. The above described argument originates from Watts and Strogatz (1998) in their seminal paper on small-worlds. In this chapter we turn this argument around by assuming that a decreasing average path-length is a structural indicator that some actors in our network have been able to form distant linkages with new partners who are not familiar to the actors existing partners. The average path-length is thus an indirect measure of distant linkages, and one could argue that it would be better to measure distant linkages directly. We argue however, that the usage of average path-length has the great advantage of only measuring 'distant linkages' that are effective in providing positional benefits to the firms involved. In order to explain why, we have visualized the ego networks of two vertices (hypothetical firms) Vi and Vj as follows:

Fig. 5.4 – Visualization of (*non effective)* distant link formation
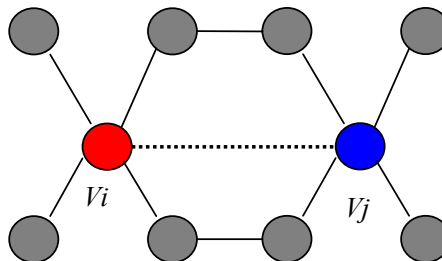
If we were to measure distant link formation 'directly', we would say that a link between two firms is a distant link, if the ego networks of both partners are non-overlapping. When using this definition the link between i and j in figure 5.4 can be defined as a distant link. However, we can also see that both firms are part of the same clique because their partners' partners are highly connected. This would make the 'distant link' between i and j much less valuable in terms of improvement of their structural network position or in terms of spreading information through the net. In figure 5.4, the dotted line between i and j represents such a 'less informative' distant

link. While *i* and *j* do not share partners, their link is not likely to bring new information to the group.

At the level of the network, the average path-length only decreases when distant linkages are formed that provide real shortcuts in the network, meaning that they connect cliques that where unconnected before. Linkages (dotted) that connect parts of a network in a way which reduces the average distance have been visualized in figure 5.5[4].

Fig. 5.5 – Visualization of *effective* distant link formation (dotted lines)
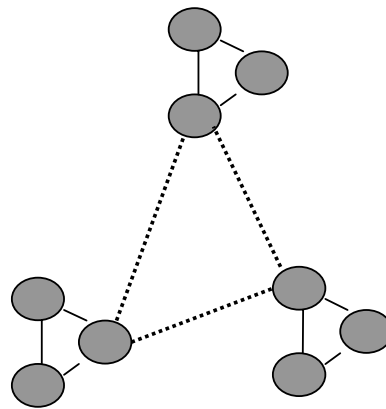
In order to understand why actors are sometimes able to create these distant linkages, and thus shorten the average distance in the network, we explore the influence of exogenous forces on the evolution of the network.

---

[4] This visualization is merely intended to clarify the difference between two ways of measuring distant link formation. The Authors are aware that this visualization cannot occur empirically because only connected actors are included in graph *N(E,V)*.

5.4 Exogenous influences on structural network change

Various studies focusing on structural network change have argued that real structural change only occurs after an exogenous shock (Barley, 1986; Piore & Sabel, 1984; Glasmeier, 1991). There are few networks studies in the organization- and management literature which take exogenous influences on network evolution into account. One important contribution comes from Madhavan, Koka, and Prescott (1998) with their study on the effect of a technological and regulatory 'event' on structural network change. Madhaven et al. (1998) find that it requires a radical technological change to enable relatively peripheral players in the network to significantly improve their network position and consequently cause a 'loosened' network structure. Koka et al. (2006) have expanded and generalized their work by providing a framework to assess the effects of exogenous events on structural network change (figure 5.1).

Exogenous or environmental changes are expressed in terms of changes in uncertainty and munificence instead of specific industry events, which makes the framework a useful meta-tool for broader applications. Following Dickson & Weaver (1997) the authors define uncertainty as "the inability of a firm's managers to accurately assess the external environment of the organization or the future changes that might occur in that environment". Uncertainty induces alliance activity (Nohria & Garcia-Pont, 1991; Powell, et al., 1996). But, whereas some authors argue that new alliances are used to reinforce a firm's relationships with existing partners during uncertainty (see Granovetter (1982) and Krackhardt (1992)), others, such as Kogut (1991), find that firms might create new alliances with new partners in order to expand its number of strategic options as a means to cope with uncertainty. To resolve the issue of either reinforcement of existing relations or the formation of new relations with new partners, the authors introduce to concept of munificence. Munificence refers to the "extend to which resources available to a firm are plentiful or scarce, after taking into account the number of firms competing for those resources" (Koka et al.2006:725). While uncertainty increases the array of actions firms can potentially make in the changing environment, the opportunities to do so are limited by the resources

available to the firm. In short, we can say that while uncertainty represents the opportunity for alliance formation with new firms, munificence represents the ability for alliance formation with new firms, given the opportunity to do so.

We use the concepts of uncertainty and munificence to hypothesize on how the genomic revolution in the beginning of the nineties has induced structural change in the pharmaceutical R&D network. In the next section, we will analyze the genomic revolution in terms of its influence on uncertainty and munificence. From there we derive our hypothesis on how we expect genomic revolution to induce relational behavior that causes structural network change.

5.4.1 The genomic revolution

The genomic revolution represents the radical scientific innovations related to the identification and understanding of the human genome and the technologies to store and analyze genetic information[5].

Our aim is to provide insight in how radical technological change such as the genomic revolution has caused structural network change. Before getting to the question of 'how', we need to defend the causal order of the question. The relational behavior of network actors could after all have created the genomic revolution and not vice versa. In order to disentangle the genomic revolution from the changes in the pharmaceutical R&D network, we show that the genomic revolution was sparked by a government led R&D program that had started long before structural changes in the network became apparent. Second, in order to be sure that government-led or government-financed R&D is not driving the structural network change, we excluded government and academic actors from the R&D network we study, thereby focusing on the industrial R&D network.

---

[5] See the website of the U.S. Department of Energy Office of Science for information on the genomic revolution and the Human Genome Project. (http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml)

Human Genome project

The Human Genome Project is a government led research project to unravel the human genetic code. The very first initiatives for the project were undertaken in 1983 by the US energy department laboratory, with the creation of DNA clone libraries representing single chromosomes. At least until 1988, the only active institutions that were involved in the setup of the human genome project were the US department of energy and the National Institute of Health. The main aims of the project were the identification of all genes and determination of sequences of chemical base pairs in human DNA, the development of storage capacity and analysis tools of genetic information, and the transference of related technologies to the private sector. The project has been completed in 2003[6].

General purpose technologies

While the Human Genome project had brought forward huge amounts of new information on genetic targets, new tools for drug discovery were needed to deal with the available genetic information. These new drug discovery tools such as combinatorial chemistry, high throughput screening and bioinformatics are not only different from conventional medicinal chemistry because they enable the testing of larger amounts of chemical entities against more drug targets, they are also much more broadly applicable in terms of disease areas and biological targets (Orsenigo, et al., 2001). Based on these new tools known as general purpose technologies (GPT), a wave of new firms specializing in GPT had been founded. In figure 5.6 we show how the proportion of alliances based on general purpose technologies has overtaken the proportion of alliances based on conventional medicinal chemistry (co-specialized technologies) in the beginning of the nineties.

---

[6] See for the timeline of developments of the Human Genome project: www.http://doegenomes.org/

Fig. 5.6 – Proportion of alliances based on General Purpose Technologies (genomics, proteomics, bioinformatics, combinatorial chemistry, high-throughput screening)
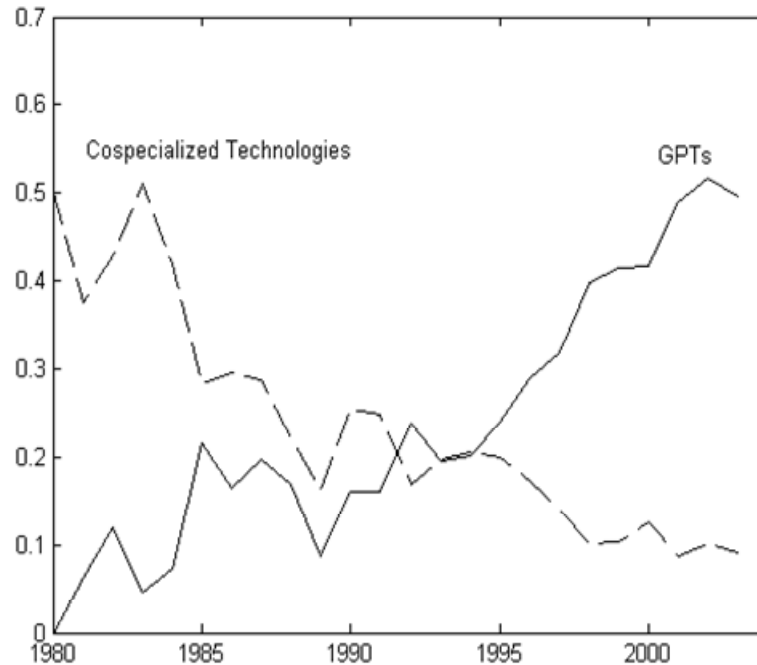
The activities of general purpose technology based firms (GPT based firms) differ from other firms' activities in drug development because they provide tools for drug development instead of developing a specific drug (Kaplan et al., 2003). According to business analysts, such as Longman (2000) and Lytton (1999), the different relational behavior of GPT firms compared to 'traditional' (co-specialized) firms, can be partly attributed to the specific characteristics of general purpose based technologies.

5.5 Effects of the radical technological change

In this section of the chapter we first investigate the effect that the genomic revolution has had on changes in munifence and uncertainty. Through these changes we hypothesize on the effects of the genomic revolution on structural network change. In the second part of the section we expand the concept of structural change by distinguishing new alliances with new 'local' partners from new alliances with new 'distant' partners. Through this approach our findings contribute to the literature on small-world networks and to the notion of the factors driving a small-world.

5.5.1 Structural change: uncertainty and munifence after the genomic revolution

The genomic revolution encompasses a number of radical scientific and technological innovations that have and are altering existing practices of drug development (Uppenbrink & Mervis, 2000; Gassmann et al.2004). While traditional approaches of medicinal chemistry and sequential experimentation have by no means become redundant, they have been complemented and intensified by general purpose technologies of which high-throughput screening, combinatorial chemistry, bio-informatics, proteomics, genomics, pharmacogenomics, and molecular design are the most important. Although these technologies are very heterogeneous in their function, together they are responsible for the alteration of drug development into a more automated, mass production process based on trial and error (Gassmann et al.2004; Nightingale, 2000; Drews, 2000). GPT are poised to improve the process of drug discovery in revolutionary ways, but there are also concerns about the increased complexity and diversity that these technologies bring to the drug development process (Longman, 2000; Orsenigo et al., 2001; Drews, 2000). Burckhardt & Brass (1990) and Hannan & Freeman (1989) find that technological change creates uncertainty, because of increased heterogeneity and complexity. Following this line of reasoning, we assume that the genomic revolution has increased environmental uncertainty.

Munificence has also increased after the genomics revolution[7]. Munificence reflects a firms' capacity (Dess & Beard, 1984). By using various new tools, general purpose technologies have greatly increased the number of possible strategies for drug discovery. This can be considered as an increase in technological resources. Adding to the notion of increasing munificence is the favorable investment climate after the Human Genome Project proved successful. Mainly general purpose based firms benefited from the willingness of investors to put their money in start-up companies that take no risks in drug development itself, but only provide the tools (Longman, 2001). The availability of technological resources together with the financial capacity to develop or invest in these resources, accounts for an increase in munificence. The increase in munificence is somewhat moderated by the wave of entrance of general purpose based firms which increases competition and lowers the average increase in munificence per firm. However, we consider firm entry to be a consequence of increasing munificence rather that a potential cause of reduced munificence. Concluding, we assume that both uncertainty and munificence have increased after the genomic revolution in the beginning of the nineties. Following the framework of Koka et al. (2006; see figure 5.1), we arrive at the following hypothesis on the effect that the genomic revolution has had on the structural changes in the pharmaceutical R&D network.

Hypothesis 1

Because the genomic revolution has increased both environmental uncertainty and munificence, we expect to see an expansion of the pharmaceutical R&D network, which we measure through a higher average portfolio size of alliances per firm in combination with an increased average range of partners per firm.

---

[7] The authors realize that while munificence clearly increased in the beginning of the nineties with the general believe that genomic based technologies would revolutionize drug development, later there has been some doubt about the revolutionizing effect of genomics. So far, drug development has mainly become more complex through all these new alternatives for development (Drews, 2000). Some analysts even say that the genomics revolution has decreased efficiency of drug development because of the decreasing number of scientists working together on one disease target. The increase in genetic targets has caused spreading of researchers over the different targets which slows down the discovery process (Longman, 2000).

5.5.2 Structural change: local clustering and distant linkages

In order to be able to hypothesize on which type of relational behavior (local link formation and/or distant link formation) to expect of firms after the genomic revolution, we first need to explain the relational behavior of firms before the genomics revolution.

Before the genomic revolution, drug development has been based on molecular biology, biochemistry, pharmacology and other disciplines for many years. Orsenigo et al. (2001) have established a connection between the nature of knowledge advancement in these years and the inter-organizational network structure in the pharmaceutical industry. They found that parallel to research in drug discovery, which develops as a branching process of older more general research hypothesis toward more specialized sub-hypothesis, a similar hierarchal branching structure unfolded in the collaborative R&D network between organizations. More specifically, the authors find that large, incumbent pharmaceutical firms manage more general knowledge in the network while new entrants (mainly dedicated biotechnology firms) specialize in specific sub-hypothesis of drug research in specialized disease areas and collaborate with the incumbent players. Over time this network evolves into a hierarchal R&D network consisting of fairly 'isolated' branches which represent specialized fields or disease areas.

After the genomic revolution, firms entered the network through unusual relational behavior. The GPT that these entering firms relied on, do not obey to the 'traditional' logic of knowledge advancement, because they contribute tools to the drug development process instead of developing drugs. With the purpose of being an aid to drug development, GPT are more broadly applied in terms of number of disease areas and biological targets. As a result, these firms mostly form non-exclusive alliances with a large variety of firms and are by definition not bound to any specific research field or disease area (Longman, 2000; Lytton, 1999). According to Orsenigo et al. (2001) GPT based firms "pertubate the structure of the network" (Orsenigo et al., 2001:490). Given the relational behavior of existing firms in the pharmaceutical R&D network in combination with the relational behavior of general purpose based firms

entering the network after the genomic revolution, we expect the network structure to be affected in the following way:

Fig. 5.7 – Simplified network topologies
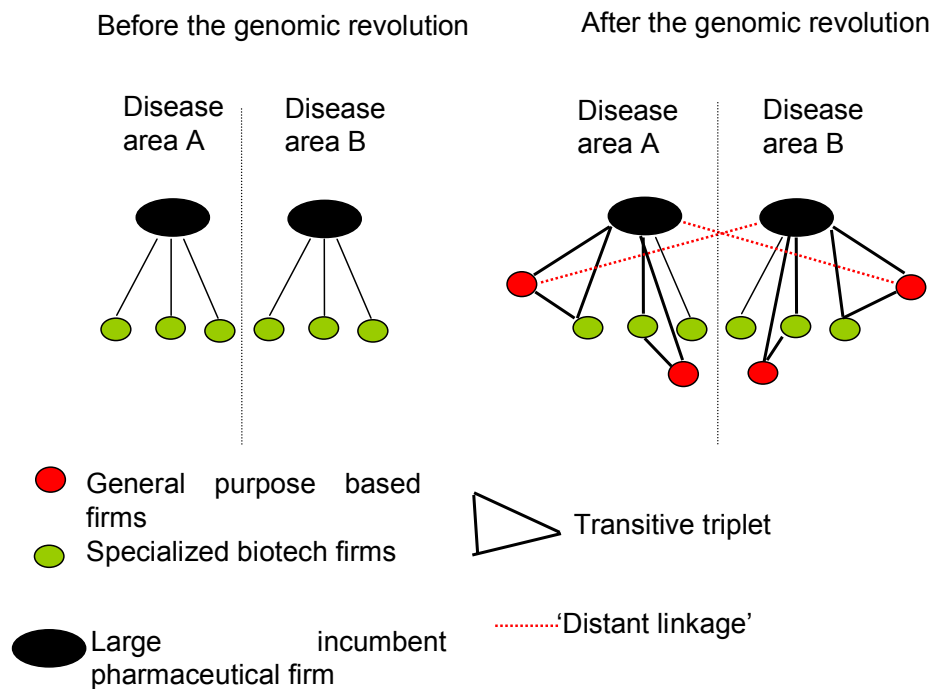
The first picture in figure 5.7 represents a simplified topology of the pharmaceutical R&D network before the genomic revolution. Within each branch (disease area A and B), specialized biotech firms collaborate with large pharmaceutical firms. It is clear that the network is organized in a hierarchal manner within isolated branches of disease areas. While this is obviously an under-representation of the complexity of the real pharmaceutical network structure, it gives us the opportunity to envision what happens when general purpose based firms enter the network with different relational behavior. The second picture in figure 5.7 shows the entrance of general purpose based firms (red nodes). The linkages they form are based on the notion that they form non-exclusive alliances with a large variety of firms and that they are by definition not bound to any specific research field or disease area. Every linkage formed by newly entering general purpose based firms, is a new linkage with a new partner, and these linkages thus cause structural change according to our definition of structural network change. The results from the new link formation of general purpose

based firms become apparent when we look at the black triangles and red (dotted) lines. The black triangles indicate that new alliances are formed locally and result in dense cliques. The red lines indicate distant linkages because they connect to distant partners from different disease areas. We can now formulate our hypothesis as follows:

Hypothesis 2 a

The genomic revolution has induced new partner search through local link formation. This relational behavior results in an increased clustering coefficient in the pharmaceutical R&D network from the beginning of the nineties.

Hypothesis 2 b

The genomic revolution has led to firms forming distant linkages between disease areas which are otherwise relatively unconnected. This relational behavior results in a decreasing average path-length of the pharmaceutical R&D network from the beginning of the nineties.

Summarizing our analytical approach in figure 5.8, we can say that exogenous change induces relational behavior that is reflected in a changed network structure. More specifically, we argue that the genomic revolution is a radical scientific innovation that is exogenous to our network. This innovation has increased environmental uncertainty and munificence for firms, who respond by increasing their portfolio size and range, leading to network expansion. In more detail, we expect that GPT based firms connect both local players within a disease area into more dense cliques (clustering coefficient), and that they connect these cliques through 'distant linkages' between disease areas (average path-length).

Fig. 5.8   –   Analytical approach

(radical) Exogenous change	Relational behavior	Structural change
Genomic revolution : general purpose technologies	Portfolio size + Portfolio range +	Network expansion
Uncertainty + Munificence +	Local link formation + Weak link formation +	Clustering coëfficient + Average Path length -

5.6 Data and methods

For our empirical analysis we have used a comprehensive and original data set that encompasses information about collaborative agreements in the biopharmaceutical industry worldwide. As a whole, the Biotech Industry Database (referred as BID) covers 20,182 collaborations subscribed by 7,407 institutions including dedicated biotech companies, established companies, specialized biotech suppliers and non-industrial research organizations since 1976. The BID has been created at the University of Siena, and was previously used by Orsenigo, Pammolli, and Riccaboni (2001) to analyze the biopharmaceutical network. As this chapter focuses on the pharmaceutical R&D network our sample consists of 10.580 collaborations among 3800 agents. For each transaction, BID includes information about:

  – Date of signing (1976-2002);
  – Partners (classified according to their role in the collaboration);
  – Stage of development at signing (i.e. discovery, preclinical, clinical…);
  – Technological content (i.e. gene therapy, genomics, combinatorial chemistry…);
  – Therapeutic category (i.e. Oncology, Metabolic disorders, Central Nervous System…);
  – Typology (viz. license, joint venture, co-development…);
  – Deal value and terms of payment (equity, upfront, milestones, royalties…).

The structure of the network of R&D collaborations can be represented by a graph *N(E,V)*, where *V* is the set of vertices (firms), and *E* are edges (R&D collaborations). Every edge *e* within the graph (industry network) is defined as a link between two partners. The graph *N* can also be represented by an *adjacency matrix N ⇔ A(N)* = [*ae*]. Matrix entry *ae* equals 1 if and only if an edge e does exist, and 0 otherwise. In order to analyze the evolution of our network over time, we took 'snapshots' of the network by labeling each connection with the date of signing. The overall graph *N(E,V)* is decomposed in time specific sub-graphs *N$_t$(E,V)*, which include all collaborations up to period *t*.

The first part of the analysis is based on simple count statistics, which reveal the changes in firm entry and alliance formation over time, to see whether the R&D network has expanded as our first hypothesis predicts. To test hypotheses 2a and 2b we calculate the clustering coefficients and average path-length respectively over time using Ucinet (Borgatti et al., 1999). Based on the adjacency matrix for each time period τ, we calculate the clustering coefficient and the average path-length. In order to facilitate interpretation of the results we calculate the clustering coefficients and average path-lengths of a random network with the same number of nodes *(V)* and the same number of linkages *(E)* at each time period τ.

5.7 Results

5.7.1 Structural change

Figure 5.9 shows the number of deal-active firms entering the network *(V)* and the number of new alliances *(E)* in the R&D network from 1967 until 2002. Deal-active firms are firms, which close at least one new alliance in period τ. Until 1992, the number of new alliances formed has been twice as high as the amount of deal-active firms entering the network in each period. After 1992 the grey line in our plot indicates that the amount of new alliances has more or less tripled the number of new deal-active firms. In other words, the average number of new alliances per firm has

increased from two to three, which means that the average size of a firms' portfolio of alliances has increased starting in the beginning of the nineties.

Fig. 5.9 – Number of new R&D collaborative agreements (E – Edges) and institutions (V – Vertices) in the network of R&D collaborations in biopharmaceuticals (1976-2002)
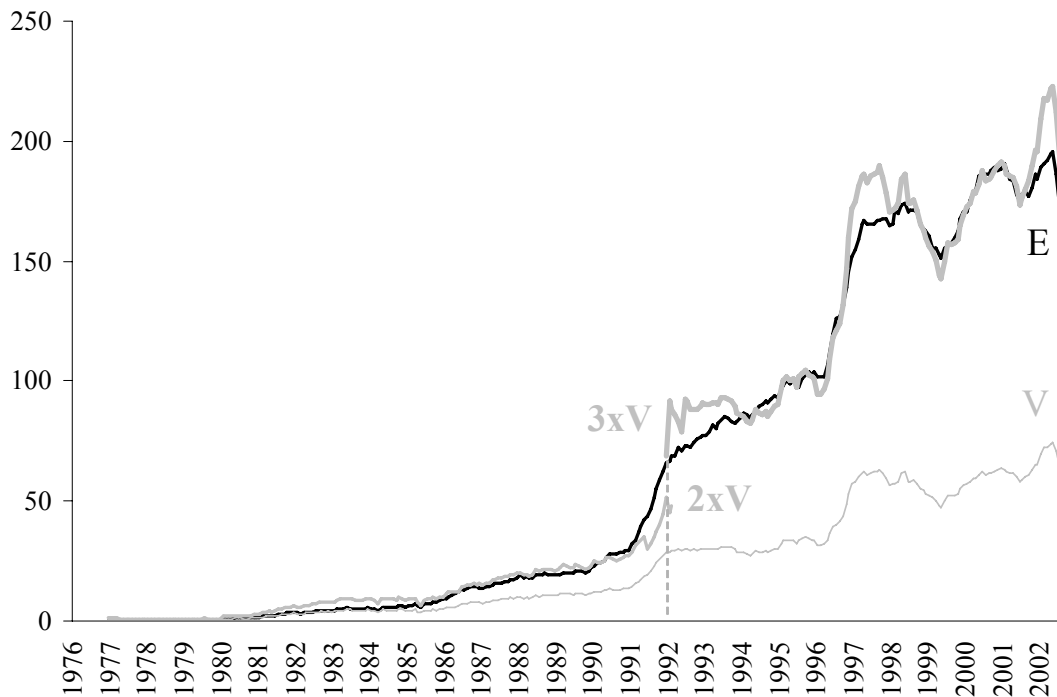
In figure 5.10 we have plotted the new alliances and new firms in the network in relation to the existing actors and their alliances. We can clearly see that there are various waves of firm entry and alliance growth over time, and the beginning of the nineties marks the start of a new wave of entry and collaboration activity.

Fig. 5.10 – Percentage change in R&D collaborative agreements (E – Edges) and institutions (V – Vertices) in the biopharmaceutical R&D network, (1976-2002)
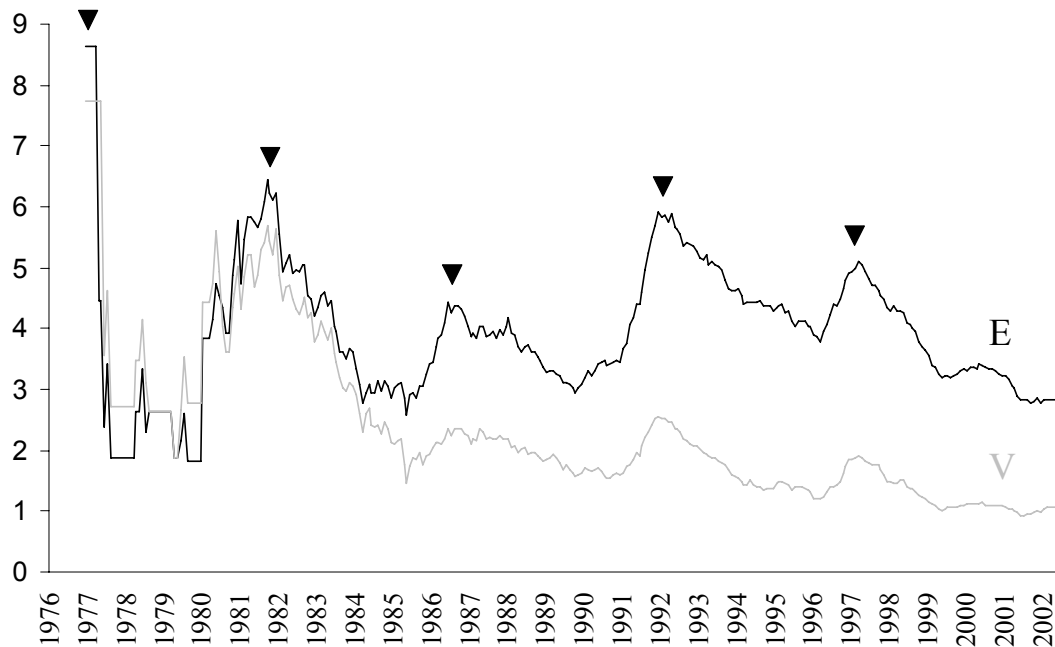
Unless newly entering firms form alliances exclusively amongst each other, a situation which is highly unlikely considering the complementary nature of their technologies, we argue that with an increased average alliance portfolio (figure 5.10), in combination with a wave of firm entrance, the average number of different partners in a firms portfolio has grown. After all, these newly entering firms are deal-active and, given the complementarities of their new technologies, they have found new partners in the existing network. In sum, we can conclude that from the beginning of the nineties, firms have on average expanded their portfolio size and range. This leads us to confirm hypothesis 1, stating that the pharmaceutical R&D network has expanded from the beginning of the nineties. From figure 5.9 and 5.10 it becomes clear however, that the average portfolio size of firms does not increase further after the beginning of the nineties and that the entrance of new firms into the network is also temporary. We can therefore conclude that although the R&D network appears to keep growing over time, the overall structural network expansion is temporary. This finding is in line with similar findings from the steel industry where technological

change leads to a temporary 'reshuffling' of relational behavior, causing temporary structural change in the network (Madhaven et al., 1998).

Local link formation

The second part of the analysis concerned a more detailed investigation of these new alliances that have been formed with new partners. We have argued that, depending on the new partner being active in 'the neighborhood' or being a 'distant' partner (or both) we expect to see different structures emerge. We start with the hypothesis of local link formation. If firms are oriented toward local link formation, we should witness an increasing clustering coefficient from the beginning of the nineties.

Fig. 5.11 – Weighted overall clustering coefficient in the pharmaceutical R&D network (full line) and of a random network (dashed line)
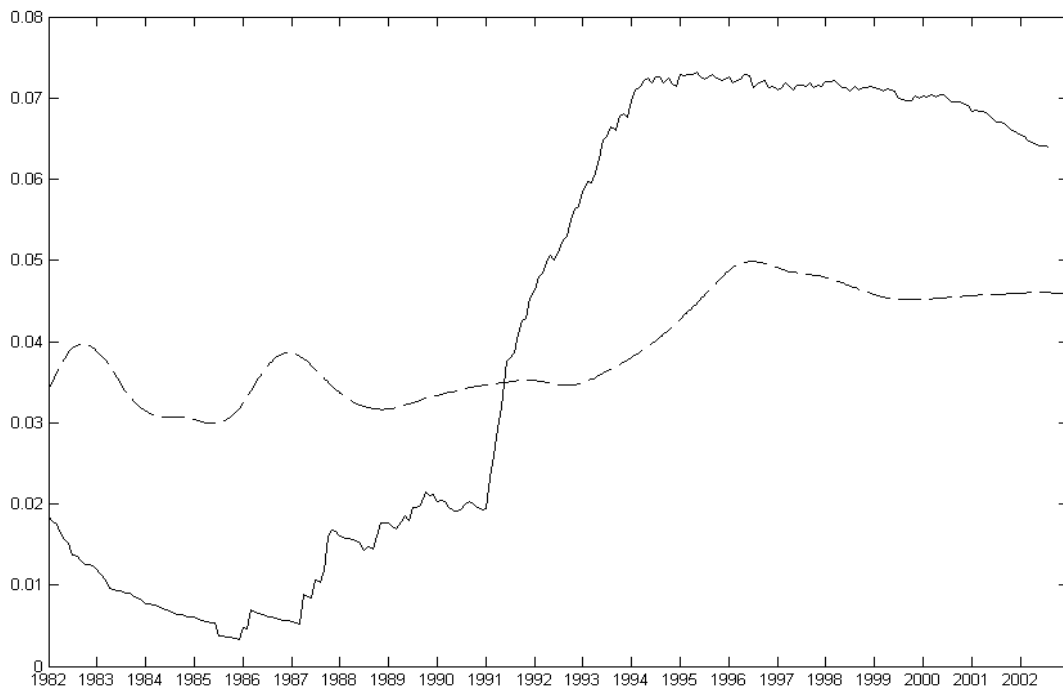
Figure 5.11 shows the weighted clustering coefficient from 1982 till 2002. Between 1991 till 1995 there is a sharp increase in the weighted clustering coefficient, while it starts to decrease slightly afterwards. This feature supports our hypothesis of an increased clustering from the beginning of the nineties, but in order to value the

increase we have added the clustering coefficient of a simulated random network (Erdös-Renyi random network). A simulated random network is a much used tool in network analysis to indicate the meaning of a certain network value. The clustering coefficient of a random network with the same amount of actors and the same average degree (portfolio size) serves as a bench-mark to compare observed relational behavior with random relational behavior. When comparing the clustering coefficient of a random network in figure 5.11 with the observed clustering coefficient in our network, the impact of the genomics revolution on local clique formation becomes even more apparent. Before the beginning of the nineties, relational behavior of firms led to less clustering compared to random relational behavior, while after the beginning of the nineties there was more clique formation then was to be expected from random relational behavior. When considering the fact that a clustering coefficient is also used in the literature as a measure of local hierarchy, our relatively low clustering coefficient before the beginning of the nineties seems to support the results of Orsenigo et al. (2001). They argue that the pharmaceutical R&D network developed as a hierarchal branching process before the nineties. Finally we can confirm hypothesis 2a through an increasing clustering coefficient after the genomics revolution, which indicates that firms have found new partners through local link formation.

Distant link formation

Did the genomic revolution bring forth (new) firms that were able to bridge some of the 'hierarchal branches' of the pharmaceutical R&D network? Given the hierarchal branching structure of the R&D network together with the inherent characteristics of GPT we expect general purpose technology based firms, who have entered the network in the beginning of the nineties, to be able to bridge these hierarchies. If so, these bridges or distant linkages would function as shortcuts in the network and shorten the average path-length. Figure 5.12 reports the average path-length of our network and the average path-length of a simulated random network over time.

  The average path-length declines from 1986, reaching its lowest point in 1995. This indicates that the reach between actors in the network has improved. After 1998 the average path-length slightly starts to increase again. At first, it seems counterintuitive

that the path-length shortens while the network grows. In our hypothesis we have argued that this shortening of the path-length has been triggered by the genomic revolution at the beginning of the nineties. Although the path-length indeed decreases, there is no indication that this decreasing trend was triggered by the genomic revolution, since the declining path-length clearly starts much earlier (from 1986). In order to value the decrease in path-length we have simulated a random network based on the exact same network size and connectivity at each time period τ. The dashed line in figure 5.12 reports the average path-length of our simulated random network. In our random network an actor has to surpass 3.5 other actors on average to reach every other reachable actor in the network. As the network grows and becomes more connected this path-length remains stable but eventually will grow according to ln(n)/ln(k) (k = average degree) for very large networks (see Watts, 1999). Starting from 1986, our empirical network and the random network slowly converge, which means that the reach between actors in the network improves despite a growing number of actors.

Fig. 5.12 – Average path-length of the pharmaceutical R&D network (full line) and of a random network (dashed line)
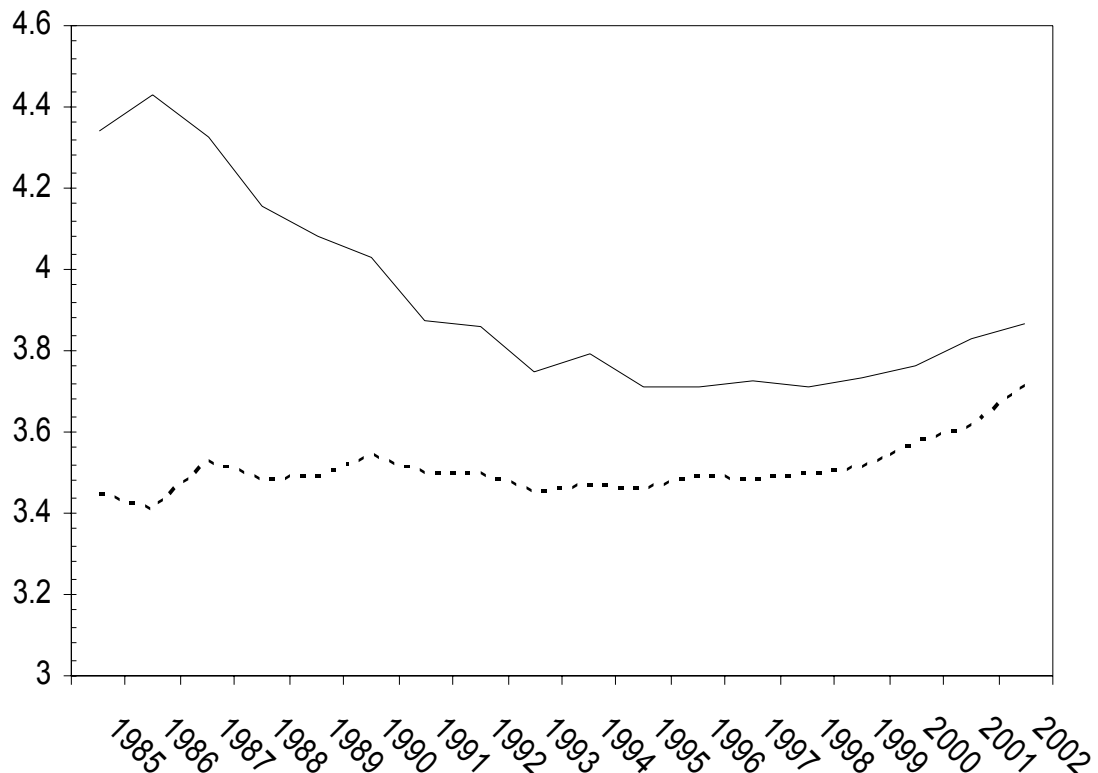
In sum, we can conclude that hypothesis 2b predicting a falling average path-length from the beginning of the nineties cannot be confirmed. However, the fact that the path-length of our random network remains relatively stable indicates that there is non-random relational behavior causing an improvement in the reach between actors.

5.8 Discussion & Interpretation

How can we interpret these results and what are its implications? To begin with we have clearly seen that the genomic revolution in the beginning of the nineties has caused structural network change. More specifically the network has expanded through an increasing average portfolio size and portfolio range, there has been a wave of firms entering and the technological focus of alliances has shifted from conventional medicinal chemistry to GPT. A second question was about the relational behavior that has caused this structural change, more specifically we studied whether firms find new partners through local linkages or whether these new partners are 'distant' partners. With this question we have extended the framework of Koka et al. (2006) on structural network change into more detail, but also we have combined the issue of partner choice in alliance networks with the study of small-world network structures. Our argument was that the relational behavior that causes a small-world structure in social networks is a combination of local link formation with relatively few distant linkages between different cliques. Together these two types of relational behavior image the combination of a regular network with a handful of random linkages, which defines a small-world according to Watts & Strogatz (1998). Thirdly, we tried to find answers to what causes this specific relational behavior which in turn causes this specific structural change. There is basic agreement among network researchers that exogenous events, which increase uncertainty, cause an increase in alliance formation, but theories differ about the partner choice following from the decision to enter an alliance. Some network researchers (e.g. Powell, 1996; Burt 1991) argue that if firms enter into new alliances with new partners, they will use their network to find these new partners, which implies structural network change to be a local growth process. This argument leaves the empirical finding of small-world

network structures unanswered however, because it fails to provide an explanation for the formation of distant linkages. We have argued that only a radical exogenous shock such as the genomic revolution can convince firms to leave their embedded path and form distant linkages with unknown partners. Following from this we hypothesized that after the genomic revolution, we would witness both local link formation and distant link formation, resembling a small-world. Although the evolution of the pharmaceutical R&D network does show a decreasing average distance and increased clustering, which are indicators of a small-world, we find no evidence for our hypothesis that the genomic revolution has induced this structural change. One explanation for this result could be that general purpose based firms, whom we expected to form distant linkages between disease branches, do not perform these alliances. The fact that we witness the increase in clustering after the genomic revolution indicates that general purpose based firms find new partners through local linkages. This latter finding confirms the structure action dynamics in network evolution, where firms choose their alliance partner using their existing network. This chapter contributes an important detail to the structure action dynamics, namely that while the underlying technological base is stable (before the genomic revolution) new alliances are formed mainly with existing partners (given the low clustering coefficient) and that it requires an exogenous radical change for firms to engage into new alliances with new partners. But even when they do, they use their network to find these partners.

Although we find no evidence for our hypothesis that the genomic revolution has caused a decreasing average distance in the network, we do witness a temporary decreasing distance between 1986 and 1998, indicating that there are firms in the network who form distant linkages that shorten the distance between other organizations. The fact that these firms already performed this relational behavior before the genomic revolution might indicate that the network structure was not composed of relatively isolated branches as we assumed, but already consisted of some organizations that improved the overall reach in the network. Jackson (2006) proposes that not only distant linkages, but also highly connected actors called 'stars' or 'hubs' can cause a path-length to decrease. While it is beyond the scope of this chapter to re-investigate the structure of the R&D network before the beginning of the

nineties, we did look for a simple indication to check Jackson's (2006) idea. Figure 5.13 plots the number of alliances for each 'type' of firm.

Fig. 5.13 – Number of alliances per firm type.

(Genomics based firms (T-DBF), established firms (EC), and co-specialized biotech (DBF))
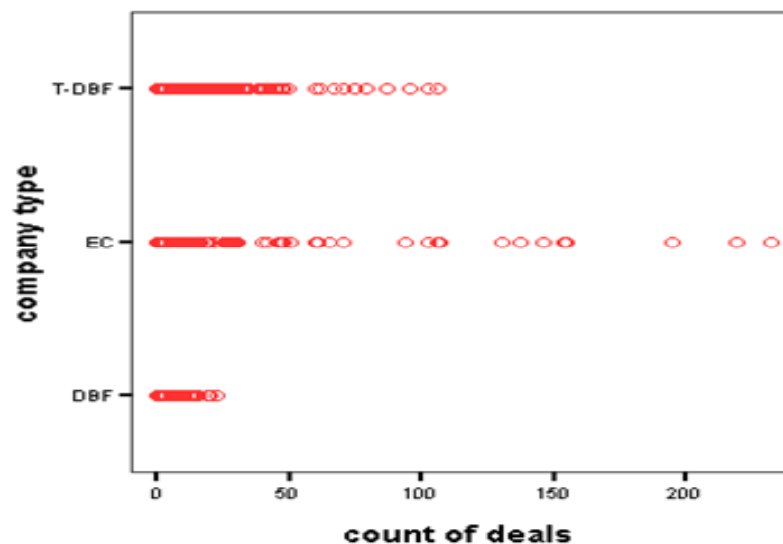
What becomes clear in this plot is that while there are quite some highly connected 'stars' amongst general purpose based firms (referred to as T-DBF), there are some pharmaceutical firms (EC) who are even more connected than general purpose based firms are. If these pharmaceutical 'stars' already connected different disease areas before the genomic revolution, then the addition of a few distant linkages by general purpose based firms would not cause a significant decrease in the average distance of the network. While being highly speculative, we feel these indications could provide an interesting start for further research.

5.9 Conclusion

The genomic revolution in the beginning of the nineties has increased environmental uncertainty and munificence, and this has led to structural changes in the pharmaceutical R&D network. The network has expanded in terms of both number of firms and number of alliances. Since alliance activity outperformed the growth in the number of firms entering the network, we can conclude that the average number of deals and the average number of different partners has increased. On a firm level this means that new alliances have been formed with new partners, causing structural network change. The formation of new alliances with new partners can take to forms, implying different structural outcomes at the level of the network. First, firms can choose their new alliance partner through their existing network, which leads to local link formation and network clustering. Second, firms can form distant linkages with unknown partners. These linkages can potentially improve a firms' position in a network if it manages to connect previously unconnected parts of the network. Taken together, local linkages and distant linkages form a small-world network which has been a popular subject of recent network studies. While we argued that a radical exogenous shock would be required for firms to be forming distant linkages, we found no evidence of this alliance behavior. We found that firms, when confronted with radical technological change keep their existing alliances and form new alliances with partners of their partners. This result is consistent with previous studies on alliance strategies and network formation. The question of what relational behavior causes a small-world in social networks remains unanswered, but some suggestions for further research are given in the discussion.

Chapter 6   Conclusion

The aim of this thesis has been to understand the strategies of firms to obtain access to relevant external knowledge that is required for innovation. Access to external knowledge has become an increasingly important part of a firm's innovation strategy. Firms are operating in markets that are more and more volatile, meaning that the underlying technologies and organizational processes are continuously changing. As a result existing products and technologies are threatened to be rendered obsolete more rapidly. Volatility induces environmental uncertainty. Firms respond to this uncertainty by building portfolios in which internal R&D and boundary spanning R&D activities are combined. By doing so, firms can exploit there existing competencies within the firm and transact with external partners to explore novel technologies without necessarily binding to them. Together, the combination of make-and-buy strategies enables a firm to flexibly adapt to a continuously changing environment[1]. However, the downside of this strategy, which is referred to as dynamic capabilities, is that external transactions are costly and risky (with respect to knowledge leakage or opportunistic behavior). But, most importantly, the knowledge that is accessed externally is not automatically usable because it can be hard to integrate with existing capabilities and to turn into new products or processes.

The decision of a firm regarding how to access external knowledge involves a trade-off between the ability to understand and integrate external knowledge on the one hand and the newness of knowledge, which is necessary to be innovative on the other hand. The newest knowledge has the highest potential for innovation, but it also entails the greatest difficulties of understanding and integrating it within the firm. By choosing a level of proximity between themselves and their external sources, firms try

---

[1] Another reason for external sourcing that has not been highlighted in the economic literature is that new entrants (in our case small biotech firms) might possess more innovative and immobile resources than established firms do. The reason for this is that in established firms, researchers are limitedly liable for their actions while researchers in start-up firms or those who found their own firms are more liable or even fully liable. As a result of this liability difference, employees with more innovative knowledge will choose full liability (including full rewards), while less competent researchers will choose to be employed by established firms where they are limitedly liable for their actions. Start-up firms thus might possess more innovative and immobile resources, which forces established firms to form strategic alliances with start-ups in order to access more innovative resources. Future co-authored work will shed more light on this issue.

to optimize the trade-off between newness and usability. This thesis is about the three dimensions of proximity which firms use to enhance their innovation capabilities: geographical proximity, cognitive proximity and relational proximity. As firms choose a level of proximity between themselves and external partners for every dimension of proximity, each firm can be positioned in a 3-dimensional action-space of geographical, relational and cognitive proximity. Moreover, this thesis addresses two main research questions:

What is the effect of different forms of proximity on a firm's ability to access relevant external knowledge?

How do these proximities affect each other in relation to a firm's ability to access relevant external knowledge?

6.1 Geographical proximity

Existing studies that address the relation between geographical proximity and innovation find that geographical proximity between organizations is mainly required to access and transfer *tacit* knowledge. (Zucker, 1996; Audretch & Feldman, 1996; Jaffe & Trajtenberg, 1993). Furthermore, tacit knowledge is found to play a more important role in exploring new knowledge that in exploiting existing capabilities. In the pharmaceutical industry the exploration of new drugs in basic research, discovery and lead optimization can indeed be said to involve much intangible tacit knowledge. Once a drug is 'discovered', testing procedures and marketing and distribution rather involve more explicit knowledge such as procedural knowledge, paper work experience and knowledge of legal issues. In Europe, almost all of the drug development activities (both explorative and exploitative) take place in a few regional clusters, while it is mainly explorative (discovery) activities that occur in the most densely clustered region surrounding London. The regional clusters surrounding London such as Cambridge, Oxford, Hertfordshire and others have emerged in the last decade as *the* dominant regions regarding biopharmaceutical activities (with 20 percent of all European organizations located in the area). Moreover, this region is characterized by a high proportion of start-up firms and a relatively strong focus on basic research in anti-cancer therapies.

While formal strategic alliances are recognized as crucial vehicles for knowledge transfer between pharmaceutical organizations, they are not being used by organizations to source knowledge locally. Moreover, alliances aimed at exploring new drugs such as R&D collaborations between universities and biotech firms are highly unlikely to occur locally. In fact, these explorative alliances are more likely to span larger geographical distances.

In sum, European biopharmaceutical organizations, especially those that are involved in exploration of new drugs, have a strong tendency to co-locate. Being co-located does however not induce knowledge transfer through formal collaboration activities among these organizations. One possible explanation is that these firms co-locate because of location-specific features such as agglomeration externalities. The emergence of the London region as the dominant region for pharmaceutical innovation activity in the last decade hints at agglomeration effects indeed. Another possible explanation could be that geographical proximity plays a more important role in informal alliances. In line with the findings of Batheld et al. (2002) our data might show only alliances that represent 'global pipelines' and not 'local buzz', as the latter consist of more informal alliances. Finally, while our findings suggest that explorative formal alliances do not require geographical proximity, it might be the case that science-based knowledge, which is known to involve relatively large amounts of tacit information, is rather accessed through other forms of proximity such as cognitive proximity or relational proximity. Geographical proximity might be relevant temporarily and thus does not necessitate co-location. Research visits or conference meetings could suffice.

6.2 Relational proximity

Relational proximity refers to how 'close' an alliance partner is within the network of inter-firm collaborations. At a distance of 1, two firms are partners, and a distance of 2 implies that two firms share a common third partner.

Relational proximity can be expressed in terms of local-, and non-local search (linkages). Local search in a network implies that alliances are formed with previous partners or with partners of previous partners. Non-local-, or distant search implies that the path of previous relations connecting two alliance partners in a network is relatively long. In this thesis we examined the long-term evolution of local and non-local alliance formation in the European and in the global network.

From 1996 until 2002, the European network of R&D collaborations is highly fragmented into small subgroups[2] where firms form only local linkages. Between 2002 and 2005 a wave of new alliances causes the clustering of these subgroups into a giant network component in which about 50 percent of all active European organizations are connected. The emergence of a giant component reveals that non-local linkages have been formed that connect previously unconnected subgroups in the network. This finding raises questions about the causes of this giant component emergence in Europe. In other words, what induces firms to leave their embedded path of existing-, and previous collaboration partners to form new alliances with unknown partners? One possible explanation has been examined in chapter 5: the effect of radical technological change. Other possible explanations are new- or abolished legislation, or a change in investment climate.

When we consider the global network of R&D collaborations from 1975 until 2002, similar structural network features become visible (as in the European network). During this period a number of radical technological changes occur, which induce changes in the alliance behavior of firms and subsequent network structures. Previous studies on alliance behavior of firms suggest that firms form mainly local alliances with partners that are already 'close' in the network of collaborations. By forming local alliances firms can benefit from being embedded in a dense local network. This embeddedness creates several advantages such as reputational effects, shared social norms, and similarity of knowledge (cognitive proximity). However, at times of radical technological change, access to radically new knowledge through distant link formation is claimed to outweigh the benefits of being locally embedded.

The genomic revolution in the beginning of the nineties can be characterized as such a radical technological change. However, even when radically new technologies that are potentially competence destroying are introduced into a market of R&D collaborators, firms are inclined to mainly collaborate within the local neighborhood of their network. At the same time, a relatively small number of firms (both large incumbents *and* new entrants) become highly central actors in the network by forming large

---

[2] One subgroup is the 'Peptido' project, which is comprised of universities and private firms that jointly work on peptides and have received considerable long-term subsidy from the European Union. An interesting future research question would be to understand what the effect of subsidies or other institutional incentives is on the emergence of a more coherent R&D network in Europe.

amounts of linkages (both local *and* non-local). These central firms (referred to as 'stars') connect previously unconnected parts of the network. As a result of the non-local linkages, the average distance between firms in the global R&D network is decreasing. A network structure that combines dense local 'neighborhoods' with distant (non-local) linkages is referred to as a 'small world'. A 'small world' network structure facilitates the diffusion of innovation in a network and prevents redundancy.

6.3 Cognitive proximity

Cognitive proximity refers to the degree of commonality in knowledge domain between two actors. The main question that we aim to answer is: to what extend is the transfer of relevant knowledge between two firms dependent on cognitive proximity between them?

While cognitive proximity has mainly been measured as a firm-level construct, more recent work has shown that it is a relational attribute, meaning that it varies with each transaction partner (Lane & Lubatkin, 1998). Cognitive proximity has therefore been analyzed at *project* level. In theory, a firm is better able to extract the rents from external transactions if it has built a sufficient stock of similar knowledge internally (Cohen & Levinthal, 1989). Through this stock of similar internal knowledge, a firm can better understand and integrate external knowledge and turn it into new products. Building on these insights, we further examined the following issues in this thesis: For which type of knowledge does cognitive proximity between projects generate innovation? How much of this 'similar' knowledge is required to generate these effects? And how does knowledge flow between internal projects and external projects in similar knowledge domains?

First of all, none of the 20 largest R&D spending global pharmaceutical firms that we investigated is performing better in research that the other firms. At the level of individual research projects, it turns out that projects are more successful if they are active in a disease area in which the firm has a critical mass of similar projects (representing high cognitive proximity). We interpret this finding as learning effects arising from cognitive proximity, although other factors could reinforce these learning effects, such as the ability of a firm to attract the best talent, and be a preferred partner for deal-opportunities outside the firm. Additionally, in-house research projects are more likely to be successful if they are active in a disease area where just a few external research projects are going on (no more than 20 percent). This finding can be interpreted as knowledge flow from external projects to internal projects. Knowledge flow occurs only when relatively few external projects are ongoing. The reason for this might be that knowledge can only be transferred from external projects to

internal projects if the same expert scientists are involved in both internal- and external projects. This would naturally limit the number of externally sourced projects. This interpretation is in line with the notion that knowledge required for pharmaceutical drug discovery is highly tacit and embedded in the scientists involved, which makes transfer of this knowledge between people leave alone projects difficult.

6.4 Interaction effects

Firms are positioned in a three dimensional action-space, which consists of geographical proximity, relational proximity and cognitive proximity. Any position of a firm in this action-space implies a level of proximity in all proximity dimensions. At the level of a regional cluster for example, firms that are geographically co-located (geographical proximity) are also more- or less connected through strategic alliances (relational proximities) and they are active in more- or less similar knowledge domains (cognitive proximity). In Europe, most biopharmaceutical organizations are geographically clustered but do not form linkages within these clusters. However, when clusters are relatively small (containing few organizations) co-located organizations are more likely to collaborate with each other, but only when they are active in a diverse set of technologies (preventing knowledge leakage). This finding is in line with the notion that geographical clusters evolve over a life cycle in which they grow to maturity. While local 'support' collaborations among interdependent organizations are valuable in the earlier stages, toward maturity non-local linkages become more valuable as organizations evolve into international players in a global network. At the level of the firm we can further conclude that firms are more likely to collaborate with partners from (or close to) their existing network, while location does not play an important role. Finally cognitive proximity appears to be an important criterion for pharmaceutical firms when choosing which research projects to invest in, both within the firm and with external partners.

Policy implications

Prevent a one-dimensional focus on geographical proximity

Policy makers often emphasize the synergy effects that arise in certain locations from the coming together of people, firms, or ideas. Policy makers should realize that while emphasizing location as the crucial element that generates synergy and innovation, implicitly they make assumptions about other forms of proximity as well. For example, our Dutch minister of education Ronald Plasterk recently argued[3] that radical innovations require the co-location of people with innovative ideas. When elaborating on this statement, he explained that these people should be specialized in different but related areas of expertise, so as to bring together complementary knowledge. In this example, the minister combines geographical proximity with cognitive proximity and assumes interaction between these people. In other words, only when the various forms of proximity between people (or firms) interact, radical innovation is claimed to occur. Numerous initiatives in the formation of science parks, creative clusters and the like are unjustly based on merely geographical proximity while the innovativeness of these regions depend for a large part on the interaction effects of multiple forms of proximity.

Creating a European small-world

An explicit aim of the European Union is to promote innovation through knowledge exchange between innovative European organizations. Our findings regarding the emergence of a small-world structure in the global network entail some insights that are useful for European policy makers that aim to stimulate diffusion of innovation through inter-organizational collaborative agreements. Our findings suggest that large pharmaceutical firms and platform technology based firms are crucial actors for European knowledge diffusion. These firms act as highly connected 'stars' that can combine local relationships with relations that connect local networks throughout the whole European network. For policy makers, the role of platform technology firms might be particularly interesting. Platform technology firms are also referred to as general purpose technology firms, which indicate that their technologies can be applied to multiple therapeutical areas. As the high degree of specialization in biopharmaceutical organizations (e.g. research labs) often forms a barrier for inter-organizational collaboration, platform technology firms can potentially break these barriers by functioning as an intermediary between otherwise unconnected organizations. Ideally, platform technology

---

[3] Ronald Plasterk in 'Zomergasten' television program. Sunday July 27.
http://www.vpro.nl/programma/zomergasten/

firms can act as brokers and stimulate cross-fertilization of knowledge between different therapeutical areas.

Thus far, we know from this thesis that the European network until 2005 is highly fragmented into local networks. Firms within these local networks are usually not co-located in a geographical sense. In order for platform technology firms to function as bridges between fragments in the network, further research is required to see whether these fragments (or local networks) are based on cognitive proximity. Only when fragments in the European network represent islands of therapeutical specialization (cognitive proximity), platform technologies can be useful to bridge these islands.

To summarize, this thesis has focused on how different forms of proximity affect a firm's ability to access relevant knowledge for innovation. We have considered the influence of being co-located in space, of being embedded in a network, and of being active in similar knowledge domains. By integrating these three proximity perspectives we contribute to various disciplines such as economic geography, organizational sociology and innovation studies. Further, this thesis investigates the make, buy or ally strategies that pharmaceutical firms employ to maximize the probability of innovation (finding new drugs). Our findings suggest that firms employ multiple governance structures simultaneously, even when targeting similar innovations. These insights contribute to our understanding of the boundaries of the firm.

While the findings in this thesis are generalizable over time and in geographical space, they are specific to the pharmaceutical industry. Nevertheless we argue that the growing importance of this industry in the future deserves a great deal of attention.

Bibliography

Abramovsky L. R. Harrison and H. Simpson (2007). 'University Research and the Location of Business R&D'. *The Economic Journal*, 117 (519) 114-141.

Adams, J. D. and A. B. Jaffe (1996). 'Bounding the effects of R&D: an investigation using matched establishment-firm data'. *The Rand Journal of Economics,* Winter (94) 700-21.

Albert, R. and A.L.Barabasi. (2002). 'Statistical mechanics of complex networks'. *Reviews of Modern Physics,* 74(47) 48-94.

Amin, A. and N. Thrift (1992). 'Neo-Marshallian nodes in global networks'. *International Journal of Urban and Regional Research,* 16(4) 571-587.

Anderson, M.L. and P. Tushman. (1986). 'Technological Discontinuities and Organizational Environments'. *Administrative Science Quarterly,* 31(3) 439-465.

Arora, A. and A. Gambardella (1990). 'Complementarity and external linkages: the strategy of large firms in biotechnology'. *Journal of Industrial Economics,* 38 (361).

Arora, A. and A. Gambardella. (1994). 'Evaluating technological information and utilizing it: scientific knowledge, technological capability, and external linkages in biotechnology'. *Journal of Economic Behavior & Organization*, 24(1) 91-114.

Arora, A. and A. Gambardella (1994). 'The changing technology of technological change: general and abstract knowledge and the division of innovative labor'. *Research Policy*, 23(5) 523-532.

Audretsch, D. B. and M. P. Feldman (1996). 'R&D Spillovers and the Geography of Innovation and Production'. *American Economic Review,* 86(3) 630-640.

Barley, S.R. (1986). 'Technology as an occasion for structuring: evidence from observations of CT scanners and the social order of radiology departments'. *Administrative Science Quarterly,* 21(1) 78-108.

Barney, J. B. (1991). 'Firm resources and sustained competitive advantage'. *Journal of Management,* 17(1) 99-120.

Batheld, H., Malmberg, A. and P Maskell (2002). 'Clusters and Knowledge: Local Buzz and Global Pipelines and the Process of Knowledge Creation'. DRUID working paper, No 02-12.

Besanko, D., Dranove, D., Shanley, M. and S. Schaefer (2007). *Economics of Strategy.* John Wiley & Sons, Chichester.

Blanc, H. and Sierra C. (1999) 'The internationalization of R&D by multinationals: a trade-off between external and internal proximity'. *Cambridge Journal of Economics,* 23 187-206.

Booth, B. and R. Zemmel (2004). 'Prospects for productivity'. *Nature Reviews Drug Discovery,* 3 451-457.

Borgatti, S., Everett, M. and L. Freeman (1999). *UCINET 5 for Windows: Software for Social Network Analysis.* Analytic Technologies, Inc., Natick, MA.

Boschma, R. (2004). 'Does geographical proximity favor innovation?' 4th Congress on Proximity Economics, Marseilles, 2-9

Boschma, R. and S. Iammarino (2007) 'Related variety and regional growth in Italy'. Working paper DRUID Summer conference, Copenhagen, June 2007.

Boschma, R.A. (2005). 'Proximity and innovation. A critical assessment'. *Regional Studies*, 39(1) 61-74.

Bourdieu, P. (1980). 'Le Capital Sociale: Notes Provisaires'. *Actes de la Recherche en Sciences Sociale,* 3 (2-3).

Breschi, S. and F. Lissoni. (2001). 'Knowledge spillovers and local innovation systems: a critical survey'. LIUC Papers in Economics 84, Cattaneo University (LIUC).

Burckhardt, M.E. and D.J. Brass (1990). 'Changing patterns or patterns of change: the effect of a change in technology on social network structure and power'. *Administrative Science Quarterly,* 35 104-127.

Burger, M., van Oort, F., and G.A. van der Knaap (2007). 'A treatise on the Scale-Dependency of Agglomeration Externalities and the Modifiable Areal unit Problem'. Working paper presented at the Kiel workshop on 'Agglomerations and Growth in Knowledge-based Societies', April 2007

Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition.* Harvard University Press, Cambridge, MA

Cassiman, B. and R. Veugelers (2006). 'In search of complementarity in the innovation strategy: Internal R&D and external knowledge acquisition'. *Management Science,* 52 (1) 68-82.

Coase, R. (1937). 'The Nature of the Firm'. *Economica,* 4(16) 386-405.

Cohen, W., and D. Levinthal (1989). 'Innovation and learning: the two faces of R&D'. *The Economic Journal,* 99 569-596.

Coleman, J.S. (1988). 'Social Capital in the Creation of Human Capital'. *American Journal of Sociology,* 94 95-120.

Cooke, P. (2001). 'Life Sciences Clusters and Regional Science Policy'. *Urban Studies,* 41 5-6.

Cowan, R. (2005). 'Network models of innovation and knowledge diffusion'. In *Clusters, Networks and Innovation*, S. Breschi and F. Malerba (eds.). Oxford University Press, Oxford, 29-53.

Deroian, P., (1986). 'On the evolution of group and network structure: Structures within structure'. *Social Networks*, 8 33-64.

Dess, G. G., and D. W. Beard (1984). 'Dimensions of organizational task environments'. *Administrative Science Quarterly*, 29 52-73.

Dickson, P., and K.M. Weaver, (1997). 'Environmental determinants and individual level moderators of alliance use'. *Academy of Management Journal*, 40 404-425.

DiMasi J. (2001). 'Risks in New Drug Development: Approval success Rates for Investigational Drugs'. *Clinical Pharmacology & Therapeutics*, May 297-307.

Dorogovtsev, S.N., and J. F. F. Mendes (2003). *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford University Press, New York.

Drews, J. (2003). 'Strategic trends in the drug industry'. *Drug Discovery Today*, 8 (9) 411-420.

Drews, J. (2000) 'Drug discovery: a historical perspective'. *Science,* 287 1960-1964.

Dussauge, P., and B. Garrette (1999). *Cooperative Strategy: Competing Successfully through Strategic Alliances*. John Wiley & Sons, Chichester

Dyer, J.H., and H. Singh (1998). 'The relational view: cooperative strategy and sources of interorganizational competitive advantage'. *Academy of Management Review,* 23 660-679.

Erickson, D. (2003). 'Wanted: drug hunters'. *In Vivo,* 45.

Feldman, M.P. (1999). 'The New Economics if Innovation, Spillovers and Agglomeration: a review of empirical studies'. *Economics of Innovation and New Technologies,* 8 5-25.

Fingleton, B. (2005). 'Testing the 'new economic geography': a comparative analysis based on EU regional data'. Working paper presented at the Kiel workshop on Trade and Location, June 2005.

Freeman, C. (1991). 'Networks of innovators: a synthesis of research issues'. *Research Policy,* 20(5) 499-514.

Gambardella, A. (1992). 'Competitive advantages from in-house scientific research: The US pharmaceutical industry in the 1980s'. *Research Policy,* 21 391-407.

Gassmann, O., Reepmeyer, G., and M. von Zedtwitz (2004). *Leading Pharmaceutical Innovation, Trends and Drivers for Growth in the Pharmaceutical Industry*. Springer, Berlin.

George, V. and A. Zaheer (2004). 'Reach Out or Reach Within? Performance Implications of Alliances and Location in Biotechnology'. *Managerial and Decision Economics,* 25 (6/7).

Glaeser, E. L., Kallal, H.D., Scheinkman, J.D., and A. Shleifer (1992). 'Growth in Cities'. *Journal of Political Economy,* 100 1126-1152.

Glasmeier A., (1991). *The High-Tech Potential: Economic Development in Rural America.* Rutgers University Press, Center for Urban Policy Research at Rutgers University.

Goyal, S., Leij, M. J. van der, and J. L. Moraga-Gonzales (2006). 'Economics: An Emerging Small World'. *Journal of Political Economy,* 114 (2) 403-412.

Granovetter, M. (1973). 'The Strength of Weak Ties'. *American Journal of Sociology*, 78 (6) 1360-1380.

Granovetter, M. (1985). 'Economic Action and Social Structure: the problem of Embeddedness'. *American Journal of Sociology,* 91 481-510.

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

Gulati, R. (1995). 'Does familiarity breed trust? The Implication of repeated ties for contractual choice in alliances'. *Academy of Management Journal*, 38 85-112.

Gulati, R., & Singh, H. (1998). 'The architecture of cooperation: Managing coordination costs and appropriation concerns in strategic alliances'. *Administrative Science Quarterly*, 43 781-814.

Gulati, R., (1998). 'Alliances and networks'. *Strategic Management Journal,* 19 293-317.

Gulati, R., and M. Gargiulo (1998). 'Where Do Interorganizational Networks Come From?' *American Journal of Sociology,* 104 (5) 439-493.

Gulati, R., Nohria, N., and A. Zaheer (2000). 'Strategic networks'. *Strategic Management Journal,* 21 203-215.

Hagedoorn, J. (2002). 'Inter-Firm R&D Partnerships: An Overview of Patterns and Trends since 1960'. *Research Policy*, 31 477-492.

Hannan, M. T., and J. Freeman (1989). *Organizational Ecology*. Harvard University Press, Cambridge (MA).

Henderson R, and I. Cockburn (1996). 'Scale, scope, and spillovers: The determinants of research productivity in drug discovery'. *RAND Journal of Economics*, 27 32-59.

Howells, J. (2002) 'Tacit Knowledge, Innovation and Economic Geography'. *Urban Studies*, 39 (5-6) 871-884.

Jackson, M.O., and B.W. Rogers (2007). 'Meeting Strangers and Friends of Friends: How Random are Socially Generated Networks?'. *American Economic Review*, 97 890-915.

Jacobs, J. (1969). *The Economy of Cities*. Random House, New York.

Jaffe, A. (1989). 'The Real Effects of Academic Research'. *American Economic Review,* 79 957-970.

Jaffe, A., and M. Trajtenberg (1993). 'Geographical Localization of Knowledge Spillovers by Patent Citations'. *Quarterly Journal of Economics,* 577-598.

Jansen, J.J.P., Bosch, F.A.J. Van den, and H.W. Volberda (2006). 'Exploratory Innovation, Exploitative Innovation, and Performance effects of organizational antecedents and environmental moderators'. *Management Science*, 52 1664.

Kaplan, S., Murray, F., and R. Henderson (2003). 'Recognition and response to biotechnology by leading pharmaceuticals companies'. *Industrial & Corporate Change,* 12 (4) 203-233.

Katila R, and G. Ahuja (2002). 'Something old, something new: A longitudinal study of search behavior and new product introductions'. *Academy of Management Journal,* 45 1183-1194.

Kogut, B., and U. Zander (1992). 'Knowledge of the firm, combinative capabilities, and the replication of technology'. *Organization Science*, 3 383-397.

Kogut, B. (1991). 'Joint ventures and the option to expand and acquire'. *Management Science,* 37 19-33.

Kogut, B. (2000). 'The Network as Knowledge: Generative Rules and the Emergence of Structure'. *Strategic Management Journal,* 21 405-425.

Koka, R. B., Madhavan, R., and J. E. Prescott (2006). 'The Evolution of Inter-firm Networks: environmental effects on patterns of network change'. *Academy of Management Review,* 31 (3) 721-737.

Krackhardt, D. (1992). 'The strength of strong ties: The importance of Philos in organizations'. In N. Nohria and R. Eccles (eds.), *Networks and Organizations: Structure, Form and Action*. Harvard Business School Press, Boston, MA.

Krugman, P. (1991). 'Increasing Returns and Economic Geography,' *Journal of Political Economy*, 99 483-499.

Lane, P. J. and M. Lubatkin (1998). 'Relative absorptive capacity and interorganizational learning'. *Strategic Management Journal,* 19 461-477.

Leiblein, M. J., and D.J. Miller (2003). 'An empirical examination of transaction- and firm-level influences on the vertical boundaries of the firm'. *Strategic Management Journal*, 24(9) 839-860.

Leiblein, M.J. (2003). 'The choice of organizational governance firm and performance: predictions from transaction costs, resource-based, and real-options theories'. *Journal of Management 29 (6)*.

Liebeskind, J. P., A. L. Oliver, A. L., Zucker, L.G., and M. Brewer (1996). 'Social Networks, Learning and Flexibility: Sourcing Scientific Knowledge in New Biotechnology Firms'. *Organization Science,* 7.

Longman, R. (2000). 'Platform Technologies and the Collaboration Paradox'. *In Vivo*, Pharmaventures (June).

Lundvall, B.A. (1992). *National Systems of Innovation: Towards a theory of innovation and interactive learning*. Pinter, London and New York.

Lytton, (1999). 'The New Structures of Platform Deals'. *Start-Up*. Pharmaventures, Oktober.

Madhavan, R., Koka, B. R., and J. E. Prescott (1998). 'Networks in transition: How industry events (re)shape interfirm relationships' *Strategic Management Journal,* 19 439-459.

March, J.G. (1991). 'Exploration and Exploitation in Organizational Learning'. *Organization Science*, 2(1) Special Issue: Organizational Learning: Papers in Honor of (and by) James G. March (1991).

Markusen, A. (1996). 'sticky places in slippery space: a typology of industrial districts'. *Economic Geography,* 27 (3) 293-313.

Mead, G. H. (1934). *Mind, self and society from the standpoint of the social behaviorist*. The University of Chicago Press, Chicago.

Milgrom, P. and J. Roberts (1990). 'The Economics of Modern Manufacturing: Technology, Strategy and Organization'. *American Economic Review,* 80 (3) 511-528.

Mowery, D.C., Oxley, J.E., and B.S. Silverman (1996). 'Strategic Alliances and Inter-firm Knowledge Transfer'. *Strategic Management Journal,* 17 77-91.

Nadler, D., and M. Tushman (1980). 'A Model for Diagnosing Organizational Behavior'. *Organizational Dynamics,* 9 (2) 35-51.

Narula, R., and G. Santangelo (2005). 'Location and R&D Alliances in the European ICT Industry' Druid Working paper No. 07-05.

Nelson, R.R., and S.G. Winter (1974). 'Neoclassical vs Evolutionary Theories of Economic Growth: Critique and Prospectus'. *Economic Journal*, 84 886-905.

Newman, M. (2001). 'The Structure of Scientific Collaboration Networks'. *Proceedings of the National Academy of Sciences,* 98 (01) 4-9.

Nightingale, P. (2000). 'Economies of Scale in Experimentation: knowledge and technology in pharmaceutical R&D'. *Industrial & Corporate Change,* 9 (2) 315-359.

Nohria, N. (1992). 'Introduction: Is the network perspective a useful way of studying organizations?'. In N. Nohria & R. Eccles (Eds.), *Networks and organizations: Structure, form and action*. Harvard University Press, Cambridge, MA.

Nohria, N., and C. Garcia-Pont (1991). 'Global strategic linkages and industry structure'. *Strategic Management Journal,* 12 105-124.

Nonaka, I. (1994). 'A dynamic theory of organizational knowledge creation'. *Organization Science*, 5 14-37.

Nonaka, I., and H Takeuchi (1995). *The Knowledge-Creating Company, How Japanese Companies Create the Dynamics of Innovation.* Oxford University Press, Oxford,

Nooteboom, B. (2004). *Inter-firm Collaboration, Learning and Networks*. Routledge, London.

O' Reilly, C. A., and M. L. Tushman (2004). 'The ambidextrous organization'. *Harvard Business Review,* (April) 74-81.

Orsenigo, L. F., Pammolli, M., and M. Riccaboni (2001). 'Technological Change and Network Dynamics: Lessons from the Pharmaceutical Industry'. *Research Policy,* 30 485-508.

Owen-Smith, J., Riccaboni, M., Pammolli. F. and W. W. Powell (2002). 'A comparison

of U.S. and European university-industry relations in the life sciences'. *Management Science,* 48 24-43.

Pammolli, F., and M. Riccaboni (2002). 'Technological Regimes and the Growth of Networks: An Empirical Analysis'. *Small Business Economics,* 19(3) 205-215.

Pammolli, F., Orsenigo, L., and M. Riccaboni (2001). 'Variety and Irreversibility in Scientific and Technological Systems: lessons from the Pharmaceutical Industry after the Molecular Biology Revolution'. In U. Pagano, A. Nicita, eds. *The Evolution of Economic Diversity*. Routledge, London.

Pavitt, K. and P, Patel (1988). 'The International Distribution of Determinants of Technological Activities'. *Oxford Review of Economic Policy,* 4 (4) 35-55.

Penrose, E. T. (1959). *The theory of the growth of the firm*. Wiley, New York.

Pfeffer, J., and G. R. Salancik, (1978). *The External Control of Organizations: A Resource Dependence Perspective*. Harper & Row, New York.

Phelps, C. (2003). 'Technological exploration: A longitudinal study of the role of recombinatory search and social capital in alliance networks'. Dissertation New York University, Graduate School of Business Administration.

Phlippen, S., and G.A. van der Knaap (2007). 'When Clusters Become Networks'. Tinbergen Institute Discussion Paper No. 2007-100/3

Phlippen, S., and M. Riccaboni (2008). 'Radical Innovation and Network Evolution'. *Annals of Economics and Statistics*, forthcoming.

Pindyck, R. S. (1991). 'Irreversibility, uncertainty, and investment'. *Journal of Economic Literature,* 29 (3) 1110-1149.

Piore, M. J. And C. F. Sabel (1984). *The second industrial divide: possibilities for*

prosperity. Basic Books, New York.

Pisano P. G. (1990). 'The R&D Boundaries of the Firm: An Empirical Analysis'. *Administrative Science Quarterly,* 35. 153-176.

Pisano, G. (1991). 'The Governance of Innovation: vertical integration and collaborative arrangements in the biotechnology industry'. *Research Policy,* 20 237-249.

Porter, M. (1990). *The Competitive Advantage of Nations*. The Free Press, New York.

Porter, M.E. (1980). *Competitive Strategy*. The Free Press, New York.

Powel W.W., Koput, K.W., and L. Smith-Doerr (1996). 'Interorganizational Innovation and the Locus of Innovation in Biotechnology'. *Administrative Science Quarterly,* 41 116-145.

Powell, W.W., Koput, K., White, D.R., and J. Owen-Smith (2005). 'Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences'. *American Journal of Sociology,* 110 (4) 1132-1205.

Prahalad, CK and G. Hamel (1990). 'The Core Competence of the Corporation'. *Harvard Business Review*, May-June, 79-91.

Prevezer, M. (1997). 'The Dynamics of Industrial Clustering in Biotechnology'. *Small Business Economics,* 9 255-271.

Rallet, A. and A. Torre (2000). 'Is geographical proximity necessary in the innovation networks in the era of global economy?' *GeoJournal,* 49 373-380.

Raub, W., and J. Weesie (1990). 'Reputation and efficiency in social interactions: An example of network effects'. *American Journal of Sociology,* 96 626-654.

Reichert, J.M. (2000). 'New biopharmaceuticals in the USA: trends in development and marketing approvals 1995–1999'. *Trends in Biotechnology,* 18 (9) 364-369.

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

Saxenian, A. (1994). *Regional Advantage*. Harvard University Press, Cambridge.

Scherer, F.M. (2007). 'Pharmaceutical Innovation'. Working paper 07-13 AEI-Brookings joint center for regulatory studies.

Schumpeter, J. A. (1942). *Capitalism, Socialism, and Democracy.* Harper and Brothers, New York.

Scott, A.J. (1993). *Technopolis: High-Technology Industry and Regional Development in Southern California*. University of California Press, Berkeley.

Senker, J., (2004). 'An overview of biotechnology innovation in Europe' in McKelvey et al (Eds.), *The Economic Dynamics of Biotechnologies*, Edward Elgar

Sidhu, J. S., Commandeur, H. J., and H. W. Volberda (2007). 'The multifaceted nature of exploration: value of supply, demand and spatial search for innovation'. *Organization Science,* 18 (1) 20-38.

Simmie J., and J. Sennet (1999). 'Innovative clusters: Global or Local Linkages?' *National Institute Economic Review,* 170 87-97.

Simon, H.A. (1955). 'A Behavioral Model of Rational Choice'. *Quarterly Journal of Economics,* 69 99-118.

Sorenson, O. and T. E. Stuart (2001). 'Syndication Networks and the Spatial Distribution of Venture Capital Investments'. *American Journal of Sociology,* 106

1546-1588.

Storper, M. (1995). 'Regional Technology Coalitions: An essential dimension of national technology policy'. *Research Policy,* 24 895-911.

Stuart, T. E., Hoang, H., and R.C. Hybels (1999). 'Interorganizational endorsements and the performance of entrepreneurial ventures'. *Administrative Science Quarterly,* 44 315-349.

Swann P, and M. Prevezer (1996). 'A comparison of the dynamics of industrial clustering in computing and biotechnology'. *Research Policy,* 25 1139-1157.

Teece, D. J., Pisano, G., and A. Shuen (1997). 'Dynamic capabilities and strategic management'. *Strategic Management Journal*, 18(7) 509-533.

Travis, J. (2008). 'Science and Commerce: science by the mass'. *Science,* 319 (5871) 1750-1752.

Tushman, M. L., and P. Anderson (1986). 'Technological discontinuities and organizational environments'. *Administrative Science Quarterly*, 31 439-465.

Uppenbrink, J., and J. Mervis (2000). 'An Information Revolution'. *Science,* 287 (5460) 1951.

Uzzi, B. (1997). 'Social structure and competition in interfirm networks: The paradox of embeddedness'. *Administrative Science Quarterly*, 42 35-67.

Uzzi, B., and J. Spiro (2005). 'Collaboration and Creativity: The Small-World Problem'. *American Journal of Sociology*, 111(2).

Van den Bosch, F. H. J, Volberda, HW., and M de Boer (1999). 'Co evolution of Firm Absorptive Capacity and Knowledge Environment: Organizational Forms and Combinative Capabilities'. *Organization Science*, 10 (5) 551-568.

Verspagen, B., and G.M. Duysters (2004). 'The small worlds of strategic technology alliances'. *Technovation*, 24 (7).

Veugelers, R. (1997). 'Internal R&D Expenditures and External Technology Sourcing'. *Research Policy*, 26 (3) 303-315.

Watts, D., and S. Strogatz (1998). 'Collective Dynamics of Small-World Networks'. *Nature,* 393 440-442.

Wilhite, A., (2001). 'Bilateral Trade and Small-World Networks'. *Computational Economics,* 18 49-64.

Williamson, O. E. (1975). *Markets and hierarchies, analysis and antitrust implications: A study in the economics of internal organization*. Free press, New York.

Wuyts, S. (2003). 'Partner Selection in Business Markets – A Structural Embeddedness Perspective' Dissertation Tinbergen Institute Rotterdam.

Zahra, S. A., and G. George (2002). 'Absorptive capacity: A review, reconceptualization, and extension'. *Academy of Management Review,* 27 (2) 185-203.

Zucker, L. G., and M. R. Darby (1996). 'Star Scientists and Institutional Transformation: Patterns of Invention and Innovation in the Formation of the Biotechnology Industry'. *Proceedings of the National Academy of Science,* 93 12709-12716.

Nederlandse samenvatting (Dutch summary)

In dit proefschrift worden de strategieën geanalyseerd die bedrijven hanteren om toegang te krijgen tot relevante externe kennis die benodigd is voor innovatie. Toegang tot externe kennis speelt een steeds belangrijkere rol in de innovatiestrategieën van het bedrijfsleven. Bedrijven opereren in markten die steeds veranderlijker worden, waarbij de onderliggende technologieën en organisatieprocessen aan voortdurende verandering onderhevig zijn. Het gevolg hiervan is dat bestaande producten en technologieën het risico lopen sneller te verouderen. Veranderlijkheid leidt tot onzekerheid over de omgeving. Bedrijven reageren op deze onzekerheid door portefeuilles samen te stellen waarin interne en externe onderzoek- en ontwikkelings activiteiten (O&O-activiteiten) worden gecombineerd. Hierdoor kunnen bedrijven gebruikmaken van de bestaande competenties die ze in huis hebben, en tegelijkertijd in samenwerking met externe partners nieuwe technologieën verkennen zonder dat ze zich hier meteen aan hoeven te binden. Deze combinatie van maak- en koopstrategieën stelt bedrijven in staat om flexibel in te spelen op een voortdurend veranderende omgeving[1]. Een nadeel van deze strategie, de zogeheten *dynamic capabilities approach*, is dat de externe samenwerking kostbaar en riskant is vanwege mogelijke kennislekkage of opportunistisch gedrag. Het belangrijkste nadeel is echter dat de door externe samenwerking gegenereerde kennis niet automatisch inzetbaar is, omdat het vaak

---

[1] Een andere reden voor externe kennisverwerving, die in de economische literatuur weinig aandacht krijgt, is het feit dat nieuwkomers (in ons geval kleine biotechnologiebedrijven) vaak over méér innovatieve en niet-mobiele medewerkers beschikken dan gevestigde bedrijven. Dit komt doordat onderzoekers in gevestigde bedrijven slechts in beperkte mate aansprakelijk zijn voor hun handelen, terwijl onderzoekers in nieuwe bedrijven en oprichters van startende bedrijven een grotere aansprakelijkheid hebben of zelfs volledig aansprakelijk zijn. Het gevolg van dit verschil in verantwoordelijkheid is dat personen met een grotere innovatieve kennis de voorkeur geven aan volledige verantwoordelijkheid (met de daarbij behorende beloning), terwijl minder competente onderzoekers eerder geneigd zijn te kiezen voor een dienstverband bij gevestigde bedrijven, waar ze slechts gedeeltelijk aansprakelijk zijn voor hun handelen. Hierdoor hebben startende bedrijven vaak méér innovatieve en niet-mobiele medewerkers in huis, zodat gevestigde bedrijven gedwongen worden strategische allianties met startende bedrijven aan te gaan om toegang te krijgen tot wetenschappers met grotere innovatieve competenties. In latere publicaties zullen we, in samenwerking met andere auteurs, meer aandacht aan dit aspect besteden.

moeilijk is deze met bestaande kennis en vaardigheden te integreren en naar nieuwe producten of processen te converteren.

Bij de beslissing van een bedrijf over de strategie om toegang tot externe kennis te krijgen, wordt een afweging gemaakt tussen enerzijds het vermogen om de externe kennis te begrijpen en te integreren, en anderzijds de nieuwheid van de kennis, die immers een voorwaarde is om innovatie te realiseren. De nieuwste kennis biedt de meeste mogelijkheden voor innovatie, maar ook de grootste problemen met begrip en integratie. Bedrijven proberen de optimale afweging tussen nieuwheid en bruikbaarheid te vertalen in een geschikte mate van nabijheid tussen het eigen bedrijf en de externe kennisbronnen. Dit proefschrift behandelt de drie dimensies van nabijheid die bedrijven hanteren om hun innovatieve vermogen te optimaliseren: geografische nabijheid, cognitieve nabijheid en relationele nabijheid. Aangezien bedrijven voor elk van deze dimensies een mate van nabijheid tussen zichzelf en hun externe partners kiezen, kan elk bedrijf worden geplaatst in een driedimensionale actieruimte van geografische, cognitieve en relationele nabijheid. Hierbij staan twee onderzoeksvragen centraal:

1   Wat is het effect van de verschillende soorten nabijheid op het vermogen van een bedrijf om toegang tot relevante externe kennis te verkrijgen?

2   Welke interacties spelen zich af tussen deze vormen van nabijheid?

6.1 Geografische nabijheid

Eerder onderzoek naar de relatie tussen geografische nabijheid en innovatie laat zien dat deze vorm van nabijheid tussen organisaties vooral van belang is voor de toegankelijkheid en overdracht van impliciete kennis (*tacit knowledge*), de kennis die is ingebed in personen, machines of organisaties (Zucker, 1996; Audretch & Feldman, 1996; Jaffe & Trajtenberg, 1993). Bovendien blijkt impliciete kennis belangrijker te zijn bij het verwerven van nieuwe kennis dan bij het gebruikmaken van bestaande competenties. In de farmaceutische industrie speelt ontastbare, impliciete kennis inderdaad een belangrijke rol in de eerste ontwikkelingsfasen van nieuwe geneesmiddelen (fundamenteel onderzoek, 'ontdekking' en lead-optimalisatie). Nadat een geneesmiddel eenmaal is 'ontdekt', berusten de testprocedures en de processen

voor marketing en distributie echter vooral op explicietere kennis, zoals kennis van procedures en wettelijke aspecten en ervaring met administratieve processen. In Europa vinden vrijwel alle activiteiten voor de ontwikkeling van geneesmiddelen (zowel de onderzoeksfase als de latere exploitatiefasen) plaats in een beperkt aantal regionale clusters, terwijl met name de onderzoeksactiviteiten (de 'ontdekking' van geneesmiddelen) zich afspelen in de regio met de sterkste concentratie: de regio rond Londen. Het afgelopen decennium hebben regionale clusters in de omgeving van Londen, zoals Cambridge, Oxford en Hertfordshire, zich ontwikkeld tot de dominante regio bij uitstek voor activiteiten op het gebied van biofarmaceutica (circa 20 procent van alle Europese organisaties is in dit gebied gevestigd). Bovendien wordt deze regio gekenmerkt door een hoog aandeel startende bedrijven en een relatief sterke nadruk op fundamenteel onderzoek naar behandelwijzen tegen kanker.

Hoewel formele strategische allianties worden erkend als essentiële factoren voor kennisoverdracht tussen farmaceutische organisaties, worden deze allianties niet door organisaties ingezet om lokaal kennis te verwerven. Bovendien zullen allianties die gericht zijn op het ontwikkelen van nieuwe geneesmiddelen, zoals O&O-samenwerkingsverbanden tussen universiteiten en biotechnologiebedrijven, in het algemeen geen lokaal karakter hebben, maar zich afspelen over grotere geografische afstanden.

Al met al hebben biofarmaceutische organisaties in Europa, met name organisaties die zich bezighouden met onderzoek naar nieuwe geneesmiddelen, sterk de neiging tot geografische concentratie (*co-location*). Deze concentratie leidt echter niet tot kennisoverdracht via formele samenwerkingsactiviteiten tussen de verschillende organisaties. Een mogelijke verklaring hiervoor is dat deze bedrijven zich in een geografische regio concentreren vanwege locatiespecifieke kenmerken, zoals externe eigenschappen van de agglomeratie. De opkomst van de regio rond Londen als de dominante regio voor innovatieve activiteiten in de farmaceutische sector in de afgelopen tien jaar, lijkt inderdaad op agglomeratie-effecten te wijzen. Een andere mogelijke verklaring is dat geografische nabijheid vooral een belangrijke rol speelt bij informele samenwerking. Mogelijk komen, in lijn met de conclusies van Batheld et al. (2002), uit onze gegevens alleen bovenregionale allianties (*global pipelines*) naar voren en niet zozeer de regionale (*local buzz*), aangezien de laatstgenoemde

betrekking hebben op meer informele samenwerkingsverbanden. Ten slotte: hoewel onze bevindingen suggereren dat formele onderzoeksgerichte allianties geen geografische nabijheid vereisen, kan het wel zo zijn dat wetenschappelijke kennis (die zoals bekend relatief veel impliciete kennis bevat) vooral toegankelijk wordt gemaakt via andere vormen van nabijheid, zoals cognitieve nabijheid of relationele nabijheid. Geografische nabijheid is vaak alleen tijdelijk van belang, zodat geografische concentratie niet noodzakelijk is. Onderlinge bezoeken door onderzoekers of bijeenkomsten op congressen zijn vaak voldoende.

6.2 Relationele nabijheid

De relationele nabijheid is een maat voor de 'afstand' tussen twee alliantiepartners binnen een netwerk van samenwerkingsverbanden tussen bedrijven. Een afstand van 1 houdt in dat twee bedrijven partners van elkaar zijn, de afstand 2 betekent dat twee bedrijven een gemeenschappelijke partner hebben, enzovoort.

Relationele nabijheid kan worden uitgedrukt in termen van lokale en niet-lokale verbanden (*linkages*). Lokale verbanden in een netwerk houden in dat er allianties worden gevormd met voormalige partners of met partners van voormalige partners. Bij niet-lokale verbanden zijn twee alliantiepartners in een netwerk met elkaar verbonden via een relatief lang pad van voormalige relaties. In dit proefschrift hebben we onderzoek verricht naar de evolutie op lange termijn van lokale en niet-lokale alliantievorming in het Europese en het mondiale netwerk.

In de periode tussen 1996 en 2002 kent het Europese netwerk van O&O-samenwerkingsverbanden een sterke fragmentatie in kleine subgroepen[2], waarbij bedrijven alleen lokale verbanden vormen. Tussen 2002 en 2005 zorgt een golf van nieuwe allianties ervoor dat deze subgroepen samenclusteren tot een kolossale netwerkcomponent waarin circa 50 procent van alle actieve Europese organisaties met

---

[2] Een van deze subgroepen is het Peptido Project, dat bestaat uit universiteiten en bedrijven die samenwerken aan de ontwikkeling van peptiden en die hiervoor een aanzienlijke langlopende subsidie van de Europese Unie hebben ontvangen. Een interessant onderwerp voor nader onderzoek is het effect dat subsidies of andere institutionele prikkels hebben op de ontwikkeling van een meer coherent O&O-netwerk in Europa.

elkaar verbonden zijn. Uit het ontstaan van deze kolossale component volgt dat er niet-lokale verbanden zijn gevormd die subgroepen op het netwerk aansluiten die voorheen niet verbonden waren. Deze conclusie werpt vragen op over de oorzaken van het ontstaan van deze kolossale component in Europa. Met andere woorden: wat drijft bedrijven ertoe hun ingebedde paden van huidige en voormalige samenwerkingspartners te verlaten om nieuwe allianties te vormen met nog onbekende partners? Een mogelijke verklaring is geanalyseerd in hoofdstuk 5: het effect van ingrijpende technologische ontwikkelingen. Andere mogelijke verklaringen zijn de invoering of afschaffing van wet- en regelgeving en veranderingen in het investeringsklimaat.

Wanneer we kijken naar het mondiale netwerk van O&O-samenwerkingsverbanden tussen 1975 en 2002, worden soortgelijke netwerkstructuren zichtbaar als in het Europese netwerk. In deze periode heeft zich een aantal ingrijpende technologische veranderingen voorgedaan, die hebben geleid tot wijzigingen in het alliantiegedrag van bedrijven en de hieruit voortvloeiende netwerkstructuren. Eerdere onderzoeken naar alliantiegedrag van bedrijven suggereren dat bedrijven vooral lokale allianties vormen met partners die zich al 'dichtbij' bevinden in het netwerk van samenwerkingsverbanden. Door lokale allianties te vormen, kunnen bedrijven profiteren van inbedding in een dicht lokaal netwerk. Deze inbedding biedt verschillende voordelen, zoals reputatie-effecten, gedeelde sociale normen en gelijkaardigheid van kennis (cognitieve nabijheid). Blijkens onderzoek weegt in tijden van ingrijpende technologische veranderingen de toegang tot radicaal nieuwe kennis via de vorming van niet-lokale verbanden echter zwaarder dan de voordelen van een lokale inbedding.

De genoomrevolutie aan het begin van de jaren 1990 kan als een dergelijke ingrijpende technologische verandering worden aangemerkt. Maar zelfs wanneer radicaal nieuwe technologieën (die in potentie kunnen leiden tot competentievernietiging) hun intrede doen op de markt van O&O-partners, hebben bedrijven de neiging om hoofdzakelijk samen te werken in de lokale omgeving van hun netwerk. Tegelijkertijd groeit een relatief klein aantal bedrijven (zowel grote, gevestigde bedrijven als startende bedrijven) uit tot belangrijke centrale actoren in het netwerk, doordat ze grote aantallen verbanden vormen (zowel lokaal als niet-lokaal).

Deze centrale bedrijven (*sterren* in het netwerk) vormen verbindingen tussen voorheen niet-verbonden segmenten van het netwerk. Ten gevolge van de niet-lokale verbanden neemt de gemiddelde afstand tussen bedrijven in het mondiale O&O-netwerk af. Een netwerkstructuur die zowel lokale, intensief verbonden segmenten (*neighborhoods*) als niet-lokale verbanden herbergt, wordt een *small world* genoemd. Een dergelijke netwerkstructuur vergemakkelijkt de verspreiding van innovatieve kennis binnen een netwerk en voorkomt redundantie.

6.3 Cognitieve nabijheid

Cognitieve nabijheid is de mate van gemeenschappelijkheid in de kennisdomeinen van twee actoren. De belangrijkste vraag die we willen beantwoorden, luidt als volgt: in hoeverre is de overdracht van relevante kennis tussen twee bedrijven afhankelijk van de cognitieve nabijheid tussen deze bedrijven?

Cognitieve nabijheid werd meestal geanalyseerd als een grootheid op het niveau van bedrijven, maar recent onderzoek heeft aangetoond dat het een relationeel attribuut is, dat mede afhangt van de transactiepartner (Lane & Lubatkin, 1998). We hebben cognitieve nabijheid daarom geanalyseerd op het niveau van projecten. In theorie kan een bedrijf een beter rendement uit externe transacties behalen wanneer het bedrijf intern een kritische massa van soortgelijke kennis heeft opgebouwd (Cohen & Levinthal, 1989). Dankzij deze interne kennis kan een bedrijf de externe kennis beter begrijpen, integreren en converteren naar nieuwe producten. Aan de hand van deze inzichten hebben we in dit proefschrift de volgende aspecten nader geanalyseerd:

- Voor welk type kennis wordt innovatie gegenereerd door cognitieve nabijheid tussen projecten?
- Hoeveel van deze 'soortgelijke' kennis is vereist om deze effecten te genereren?
- Hoe verlopen de kennisstromen tussen de interne en externe projecten in soortgelijke kennisdomeinen?

Om te beginnen: geen van de door ons onderzochte 20 farmaceutische multinationals met de hoogste O&O-uitgaven presteert op onderzoeksgebied beter dan de andere

bedrijven. Op het niveau van individuele onderzoeksprojecten blijkt dat projecten succesvoller zijn indien ze betrekking hebben op een ziektegebied waarop het bedrijf een kritische massa van soortgelijke projecten heeft opgebouwd (en daarmee een grote cognitieve nabijheid). We verklaren dit resultaat uit leereffecten die voortvloeien uit cognitieve nabijheid, hoewel ook andere factoren deze leereffecten kunnen versterken, zoals het vermogen van een bedrijf om de beste talenten aan te trekken of om een gewilde partner te zijn voor potentiële samenwerkingsverbanden met derden. Bovendien zijn interne onderzoeksprojecten vaker succesvol als ze betrekking hebben op een ziektegebied waarop slechts weinig externe onderzoeksprojecten gaande zijn (niet meer dan 20 procent). Dit resultaat kan worden verklaard door een kennisstroom van externe naar interne projecten. Deze kennisstroom treedt alleen op als er relatief weinig externe projecten gaande zijn. Een mogelijke oorzaak hiervoor is dat kennis alleen van externe naar interne projecten kan worden overgedragen indien dezelfde wetenschappers betrokken zijn bij zowel de interne als de externe projecten. Dit houdt uiteraard een praktische beperking in voor het mogelijke aantal projecten met externe kennisverwerving. Deze interpretatie sluit aan op de gedachte dat de voor het ontdekken van geneesmiddelen vereiste kennis een sterk impliciet karakter heeft en in de betrokken wetenschappers is ingebed, waardoor de overdracht van deze kennis tussen personen en vooral ook tussen projecten wordt bemoeilijkt.

6.4 Interactie-effecten

Bedrijven kunnen worden gepositioneerd in een driedimensionale actieruimte, gekenmerkt door de dimensies van geografische, relationele en cognitieve nabijheid. De positie van een bedrijf in deze ruimte bepaalt een mate van nabijheid in elk van deze drie dimensies. Op het niveau van een regionaal cluster geldt bijvoorbeeld dat bedrijven die in een bepaalde regio geconcentreerd zijn (geografische nabijheid), ook in meerdere of mindere mate verbonden worden door strategische allianties (relationele nabijheid) en actief zijn in min of meer soortgelijke kennisdomeinen (cognitieve nabijheid). In Europa zijn de meeste biofarmaceutische organisaties geografisch geclusterd, maar ze vormen geen verbanden binnen deze clusters. Wanneer clusters echter relatief klein zijn (met een gering aantal organisaties), is de

kans groter dat bedrijven uit dezelfde regio met elkaar samenwerken, maar alleen wanneer ze op uiteenlopende technologische gebieden actief zijn (zodat kennislekkage wordt verhinderd). Deze conclusie sluit aan op de gedachte dat geografische clusters binnen een zekere levenscyclus tot rijping komen. Terwijl lokale, 'ondersteunende' samenwerking tussen gerelateerde organisaties in de beginstadia waardevol is, worden tijdens de volwassenwording van het cluster de niet-lokale verbanden belangrijker, wanneer de organisaties zich ontwikkelen tot internationale spelers in een wereldwijd netwerk. Op bedrijfsniveau kunnen we concluderen dat bedrijven eerder geneigd zijn tot samenwerking met partners uit (of op korte afstand van) hun huidige netwerk, waarbij geografische locatie geen belangrijke rol speelt. Ten slotte blijkt cognitieve nabijheid voor farmaceutische bedrijven een belangrijk criterium te zijn bij de selectie van onderzoeksprojecten om in te investeren, zowel binnen de eigen organisatie als met externe partners.

Implicaties voor het beleid

Eendimensionale focus op geografische nabijheid vermijden

Beleidsmakers benadrukken vaak de synergie-effecten die op bepaalde locaties ontstaan uit het bijeenkomen van mensen, bedrijven en/of ideeën. Hierbij dienen ze zich echter te realiseren dat ze door de nadruk te leggen op locatie als doorslaggevend element voor het genereren van synergie en innovatie, impliciet ook aannames maken over andere vormen van nabijheid. Ronald Plasterk, de Nederlandse minister van onderwijs, heeft onlangs betoogd[3] dat het voor radicale innovaties noodzakelijk is dat er *co-location* plaatsvindt van mensen met innovatieve ideeën. Hij beschreef vervolgens dat deze mensen gespecialiseerd dienen te zijn in verschillende, maar onderling gerelateerde expertisegebieden, zodat er complementaire kennis bijeen wordt gebracht. In dit voorbeeld combineert de minister geografische nabijheid met cognitieve nabijheid en veronderstelt hij een interactie tussen deze mensen. Met andere woorden: radicale innovatie zal alleen optreden als er interactie plaatsvindt tussen verschillende vormen van nabijheid tussen mensen (of bedrijven). Veel initiatieven voor de vorming van *science parks*, creatieve clusters en dergelijke zijn ten onrechte uitsluitend gebaseerd op geografische nabijheid, terwijl de innovatieve kwaliteit van deze regio's voor een groot deel afhankelijk is van de effecten van de interactie tussen de verschillende soorten nabijheid.

Een Europese 'small world'

Een expliciete doelstelling van de Europese Unie is het bevorderen van innovatie door kennisuitwisseling tussen innovatieve Europese organisaties. Onze conclusies over het ontstaan van een *small world*-structuur in het mondiale netwerk bieden een aantal inzichten die waardevol zijn voor Europese beleidsmakers die streven naar diffusie van innovatieve kennis door middel van samenwerkingsovereenkomsten tussen organisaties. Uit onze conclusies komt naar voren dat grote farmaceutische bedrijven

---

[3] Ronald Plasterk in het televisieprogramma *Zomergasten*, 27 juli 2008. http://www.vpro.nl/programma/zomergasten/

en leveranciers van platformtechnologieën een cruciale rol spelen bij de diffusie van kennis in Europa. Deze bedrijven fungeren als sterk verbonden 'sterren' in het netwerk, die lokale relaties kunnen combineren met relaties tussen lokale netwerken binnen het Europese netwerk als geheel. Voor beleidsmakers is de rol van platformtechnologieleveranciers waarschijnlijk van extra belang. Leveranciers van platformtechnologieën worden ook wel algemeen-technologische bedrijven genoemd, omdat hun technologieën kunnen worden toegepast op verschillende therapeutische deelgebieden. De sterke mate van specialisatie bij biofarmaceutische organisaties (denk aan onderzoekslaboratoria) vormt vaak een barrière voor samenwerking tussen organisaties, maar platformtechnologieleveranciers kunnen deze barrières helpen doorbreken door te fungeren als schakel tussen organisaties die verder geen verbanden hebben. In het ideale geval kunnen platformtechnologieleveranciers als tussenpersonen optreden en de kruisbestuiving van kennis tussen de verschillende therapeutische deelgebieden bevorderen.

We weten uit dit proefschrift dat het Europese netwerk tot 2005 sterk is gefragmenteerd in lokale netwerken. Bedrijven binnen deze lokale netwerken zijn doorgaans niet geografisch geconcentreerd. Om platformtechnologieleveranciers een brugfunctie tussen de netwerkfragmenten te kunnen laten vervullen, is nader onderzoek nodig om te zien of deze fragmenten (of lokale netwerken) zijn gebaseerd op cognitieve nabijheid. Alleen indien de fragmenten in het Europese netwerk samenvallen met eilanden van therapeutische specialisatie (cognitieve nabijheid), kunnen platformtechnologieën een nuttige brugfunctie tussen deze eilanden vervullen.

Samengevat kunnen we zeggen dat dit proefschrift aandacht besteedt aan de invloed van verschillende vormen van nabijheid tussen bedrijven op het vermogen van bedrijven om toegang te krijgen tot kennis die nodig is voor innovatie. We hebben gekeken naar de invloed van (1) geografische concentratie, (2) inbedding in een relatienetwerk en (3) actief zijn in soortgelijke kennisdomeinen. Door integratie van deze drie nabijheidsperspectieven kunnen we een bijdrage leveren aan verschillende vakgebieden, zoals economische geografie, organisatorische sociologie en innovatie studies. Verder wordt in dit proefschrift onderzoek gedaan naar de maak-, koop- of alliantie-strategieën die farmaceutische bedrijven hanteren om de waarschijnlijkheid

van innovatie (het ontdekken van nieuwe geneesmiddelen) te maximaliseren. Uit onze resultaten komt naar voren dat bedrijven gelijktijdig meerdere beheersstructuren toepassen, zelfs bij het nastreven van soortgelijke innovaties. Deze inzichten verhelderen ons inzicht in de grenzen van de onderneming.

Hoewel de conclusies in dit proefschrift over de tijd en de geografische ruimte gegeneraliseerd kunnen worden, zijn ze wel specifiek voor de farmaceutische industrie.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. The following books recently appeared in the Tinbergen Institute Research Series:

384. M.J. VAN DER LEIJ, *The economics of networks: Theory and empirics.*
385. R. VAN DER NOLL, *Essays on internet and information economics.*
386. V. PANCHENKO; *Nonparametric methods in economics and finance: dependence, causality and prediction.*
387. C.A.S.P. SÁ, *Higher education choice in The Netherlands: The economics of where to go.*
388. J. DELFGAAUW, *Wonderful and woeful work: Incentives, selection, turnover, and workers' motivation.*
389. G. DEBREZION, *Railway impacts on real estate prices.*
390. A.V. HARDIYANTO, *Time series studies on Indonesian rupiah/USD rate 1995 – 2005.*
391. M.I.S.H. MUNANDAR, *Essays on economic integration.*
392. K.G. BERDEN, *On technology, uncertainty and economic growth.*
393. G. VAN DE KUILEN, *The economic measurement of psychological risk attitudes.*
394. E.A. MOOI, *Inter-organizational cooperation, conflict, and change.*
395. A. LLENA NOZAL, *On the dynamics of health, work and socioeconomic status.*
396. P.D.E. DINDO, *Bounded rationality and heterogeneity in economic dynamic models.*
397. D.F. SCHRAGER, *Essays on asset liability modeling.*
398. R. HUANG, *Three essays on the effects of banking regulations.*
399. C.M. VAN MOURIK, G*lobalisation and the role of financial accounting information in Japan.*
400. S.M.S.N. MAXIMIANO, *Essays in organizational economics.*
401. W. JANSSENS, *Social capital and cooperation: An impact evaluation of a women's empowerment programme in rural India.*
402. J. VAN DER SLUIS, *Successful entrepreneurship and human capital.*
403. S. DOMINGUEZ MARTINEZ, *Decision making with asymmetric information.*
404. H. SUNARTO, *Understanding the role of bank relationships, relationship marketing, and organizational learning in the performance of people's credit bank.*
405. M.Â. DOS REIS PORTELA, *Four essays on education, growth and labour economics.*
406. S.S. FICCO, *Essays on imperfect information-processing in economics.*
407. P.J.P.M. VERSIJP, *Advances in the use of stochastic dominance in asset pricing.*
408. M.R. WILDENBEEST, *Consumer search and oligopolistic pricing: A theoretical and empirical inquiry.*