When Clusters become Networks

*Sandra Phlippen**

Bert van der Knaap

Department of Applied Economics, Erasmus University Rotterdam

* Tinbergen Institute

When clusters become networks

A study into the causes of strategic collaboration amongst geographically clustered organizations[1]

Sandra Phlippen[2]

Tinbergen Institute

Erasmus University Rotterdam

phlippen@few.eur.nl

Bert van der Knaap

Department of Applied Economics

Erasmus University Rotterdam

vanderknaap@few.eur.nl

Abstract

Policy makers spend large amounts of public resources on the foundation of science parks and other forms of geographically clustered business activities, in order to stimulate regional innovation. Underlying the relation between clusters and innovation is the assumption that co-located firms engaged in innovative activities benefit from knowledge that diffuses locally. In order to access this knowledge, firms are often required to form more- or less formal relations with co-located firms. Empirical evidence shows however that besides some success cases like Silicon Valley and the Emilia- Romagna region where firms collaborate intensively, many regional clusters are mere co-locations of firms. To enhance our understanding of why some clusters become networks of strategic collaboration and others don't, we study link formation within European biopharmaceutical clusters. More specifically we look at the effect of cluster characteristics such as number of start-up firms, established firms or academic institutions, or the nature of the collaborations on the probability of local link formation

Keywords: regional clusters; networks; local & global linkages; pharmaceutical industry

JEL codes: R11, R12, R58, D83, D85

---

[2] Corresponding author.

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 paper 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 paper 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. Our paper 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 paper will discuss the data and research methodology, followed by the results in section 5. In the concluding section (section 6) of the paper 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).

2 Preliminary data view

In figure 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.
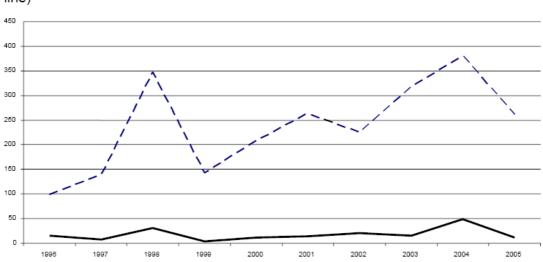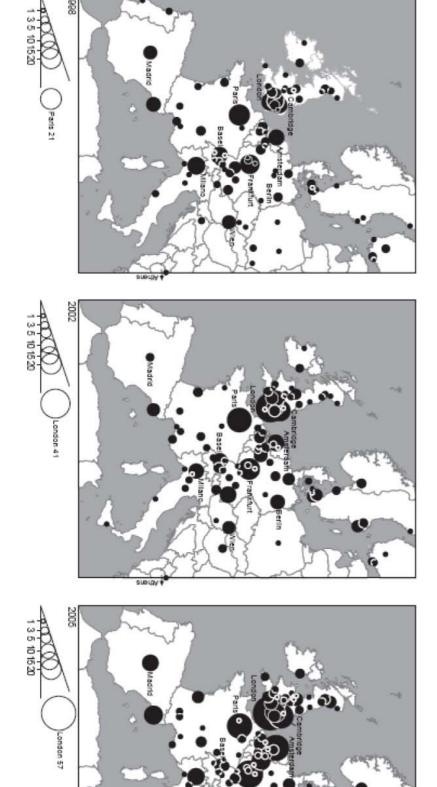
Fig.1 – Annual new collaborative non-local agreements (dotted line) and local agreements (full line)

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 1 and figures 2, 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. 1 – European pharmaceutical clusters in 1999, 2002 and 2005 (descriptive)

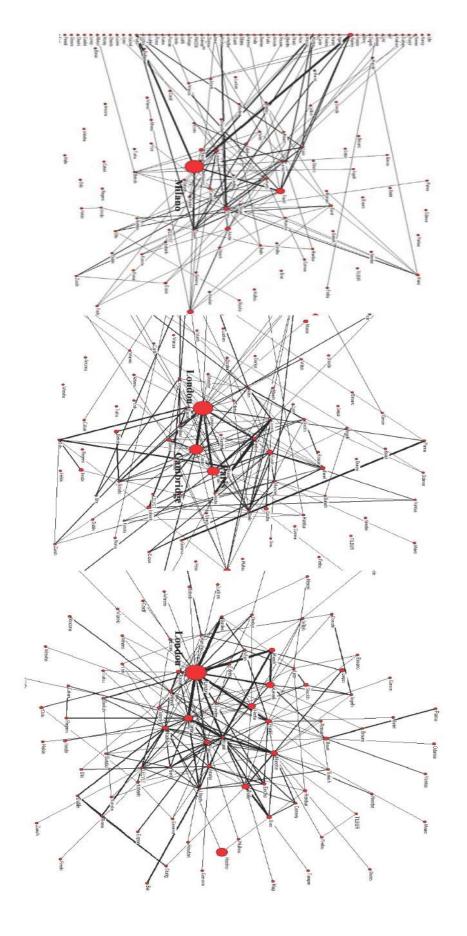|	June 1999	June 2002	June 2005
Mean number of organizations per cluster	5.79	7.32	10.04
Stand dev. Of organizations per cluster	4.96	7.73	11.10
Linkages	687	715	1157
Organizations	303	411	609

Fig.2 – European Pharmaceutical clusters in 1999, 2002 and 2005. Node size represents number of deal-active firms

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 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 1 and Figure 3, it is no surprise that alongside the increasing number of deal-active organizations, the number of alliances has increased as well.

Fig. 3 - Collaboration between clusters (ties) & within clusters (node size) in 1999, 2002 and 2005

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 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 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 focusses 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.

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.

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.

Local link formation

3.1 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 many upstream alliance activities (focusing on exploration) have a higher probability of local collaboration

3.2 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 paper, 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.

3.3 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.

3.4 (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 cause 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. We expect to find an inversed u-shaped relation between the number of deal active organizations and the probability to collaborate locally. This relation would indicate decreasing returns of local collaboration.

4 Data and methods

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 2 summarizes the terminology and definitions of our clusters.

Table 2 – Cluster definitions and characteristics.

- Cluster name = the city name of a cluster where most companies are located
- Size = number of deal active organizations located in the cluster
- Alliances = number of alliances closed by organizations located in the cluster
  - Intra-cluster alliances = number of alliances where both partners where located within the cluster
  - Inter-cluster alliances = number of alliances with a partner located outside the cluster
- Company types per cluster = number of start-up firms; academic organizations; financial firms; established firms; Government organizations
- Therapeutic 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 product development process of alliances formed by organizations located in a cluster (discovery, preclinical, clinical 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 3) below.

Table 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

4.2 Dependent variable

As our aim in this paper 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.

4.3 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 paper, 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 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 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.

5 Results

Table 4 shows the regression results, which are presented separately for each hypothesis.

Table 4 - Negative binomial regression results using maximum likelihood estimations

Dependent Variable: local link formation				Dependent Variable: non-local link formatic			
Coefficient	Std. Error	Z-Statistic	Prob.	Coefficient	Std. Error	z - Statistic	Prob.
Effect of upstream & 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				

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.

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.

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 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'.

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 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.

---

[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.

VI Discussion & conclusion

The primary aim of this paper 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 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). 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 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.

References

Abramovsky, L., Harrisson R., and Simpson H. (2007) "University Research and the Location of Business R&D" Economic Journal 117: 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.

Arora, A. and A. Gambardella (1990). "Complementarity and External Linkages: The Strategies of Large Firms in Biotechnology." The Journal of Industrial Economics 38:361-379.

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.

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.

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, Vol. 23, pp. 187-206.

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

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

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

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

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

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., H. D. Kallal, J. D. Scheinkman, and A. Shleifer. (1992). "Growth in Cities." Journal of Political Economy 100:1126-1152.

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

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

Gulati, R. (1998) "Alliances and Networks". Strategic Management Journal, Vol. 19, pp. 293-317

Hagedoorn, J. (2002) 'Inter-Firm R&D Partnerships: An Overview of Patterns and Trends since 1960', Research Policy, vol. 31, pp. 477-92

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

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

Jaffe, A., M. Trajtenberg (1993). "Geographical Localization of Knowledge Spillovers by Patent Citations." Quarterly Journal of Economics: 577-98.

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

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

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

Markusen, A. (1996) "Sticky Places in Slippery Space: a Typology of industrial Districts" Economic Geography 72: 293-313.

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

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

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

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.

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.

Piore, M. J. And C. F. Sabel (1984). The second industrial divide: possibilities for prosperity. New York, Basic Books.

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

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.

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.

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

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

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

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.

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

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

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 (November): 12709-12716.

Zucker, L.G., Darby, M.R., and J.S. Armstrong (2002) "Commercializing Knowledge: University Science, Knowledge Capture and Firm performance in Biotechnology" Management Science 48: 138-153.

Appendix I

Scatter plot of cluster size (nodes) and local collaboration (local coll_y)