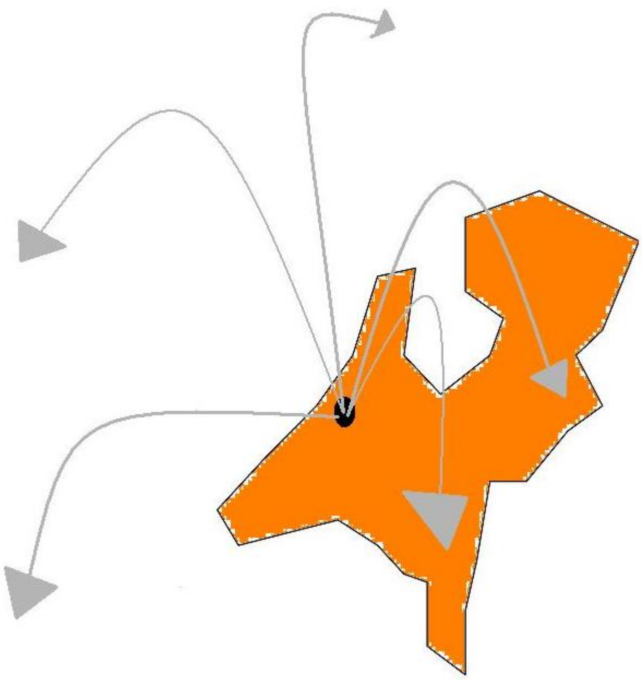
ERASMUS UNIVERSITY ROTTERDAM

ERASMUS SCHOOL OF ECONOMICS

Economics and Business

"KNOWLEDGE DIFFUSION AND LABOR MOBILITY: A LEIDEN BIO SCIENCE PARK CASE STUDY"



Master thesis

Supervisor: S. Phlippen

Date: 16th December 2010

Author: Jelte Dijkstra

E-mail: Dijkstra.Jelte@gmail.com

ABSTRACT

Knowledge is the foundation for innovation and plays a major role in the current economic environment. Especially in high tech industries such as the biotechnology sector, the role of knowledge has become increasingly important and has played a role in the formation of numerous clusters worldwide. There is a heated debate about how and where knowledge diffuses. This thesis is constructed as a case study in which the diffusion of knowledge and one of the suggested knowledge diffusion mechanisms are investigated. Knowledge created on the Leiden Bio Science Park (LBSP), a typical biotech cluster located in Leiden, The Netherlands provides the research setting. The nature of the thesis is for a large part descriptive but the goal is to test to what extent the knowledge diffuses locally and how large the role of labor mobility in this knowledge diffusion. By following the trail of patent citations of a sample of LBSP patents, the knowledge diffusion can be determined. A new data resource (LinkedIn) provides more detailed information about the labor mobility of the LBSP inventors, which leads to more detailed insights in their actual job mobility. The knowledge diffusion doesn't seem to diffuse locally. The role of labor mobility as a knowledge transfer mechanism for LBSP knowledge seems to be negatively related to the distance the knowledge is transferred.

KEYWORDS; Leiden Bio Science Park, biotechnology, clusters, knowledge diffusion, knowledge spillovers, labor mobility.

ACKNOWLEDGEMENTS

You always hear people complain about the difficulties they encounter when writing their thesis, well, I can assure you they are all true ;)

The thesis marks the end of my student era. Therefore I would like to express my gratitude to all who contributed in various ways to the realization of my graduation in general and this thesis specifically.

First of all, I would like to thank my supervisor, Sandra Phlippen, for her guidance throughout the process, passing on her enthusiasm and most of all for her patience.

Furthermore I would like to thank my parents for their moral support and confidence, Anouk for her patience and Petit Paris for the great coffee and inspiring moments.

Amsterdam, 16th December 2010,

Jelte Dijkstra

INDEX

INTRODUCTION

On the 16th of December of 1990 the Leiden Bio Science Park (LBSP) caught the world's attention with the birth of bull Herman, at that time the world's first genetically modified bull. The LBSP is a typical example of a high tech, knowledge based cluster which are described by Porter (1990) to *"dominate the current economic map"*. Clustering or co-locating is the spatial aggregation of a certain type of economic activity (Porter, 1998). Operating in a cluster is described to increase the organizations' productivity and innovative ability. The LBSP cluster, located in Leiden, The Netherlands, plays a central role in this thesis.

For organizations residing in high tech, knowledge based clusters such as the LBSP, the ability to come up with new and better ways to raise quality and improve productivity growth is determining their long-term ability to prosper (Porter 1990; Nonaka, 1994; Grant 1996 in Malmberg and Power, 2005). Innovative ability is not only a major competitive competence in the LBSP biopharmaceutical industry, but for all high tech and knowledge based industries. Knowledge is the foundation of innovation and the inducement for this thesis. Gaining further insight in how organizations can gain access to knowledge, the actual diffusion of knowledge and mechanisms that play a role in the diffusion are therefore of great societal and economic value and are the goals of this thesis.

The fact that knowledge diffuses is widely acknowledged in the literature, but there is a heated debate going on about where knowledge diffuses to and what facilitates the diffusion. There are a lot of researchers claiming that knowledge shows a vast tendency to flow locally (Jaffe et al., 1993; Audretsch and Feldman, 1996). The availability of these local knowledge flows are therefore an important motivation for organizations to allocate inside a cluster and for the authorities to invest public money in the development of these clusters. At the same time different mechanisms are pointed out to facilitate the transfer of knowledge, even though researchers have not found consensus about what these facilitators are. There exist operational difficulties that make it troublesome to track knowledge flows and identify the facilitating mechanisms.

The importance of knowledge in the current economic environment and the ongoing ambiguity on the knowledge diffusion mechanisms have inspired me to write about this subject. In this thesis I address these topics of knowledge diffusion and knowledge diffusion mechanisms by means of taking the LBSP as a case study. I investigate the knowledge diffusion of knowledge developed and patented by organizations residing on the LBSP. After having determined the knowledge flows, I focus on one of the suggested diffusion mechanisms: labor mobility. By tracking the job mobility of the inventors that have developed the designated LBSP patents, I find out to what extent this diffusion mechanism plays a role. By applying a new method of data collection I dig up more complete data on the mobility of the inventors. Thereby I exactly monitor the impact of labor mobility in the diffusion of LBSP knowledge.

This brings us to the research question of this thesis:

Where does knowledge generated on the Leiden Bioscience Park flow to and to which extent does labor mobility play a role in facilitating these knowledge flows?

The thesis is build up of the following parts. Chapter 2 is the start of the theoretical framework. It describes the role of knowledge and innovation in the biotech sector and elaborates on the concept of knowledge. Chapter 3 describes the requirements for an effective knowledge transfer. Chapter 4 elaborates on the knowledge transfer mechanisms in general and focuses on the mechanism of labor mobility. Chapter 5 presents the methodology. Data source, data collection and methods used are treated in this chapter. Chapter 6 handles the results, descriptive statistics and analysis where the research question is answered. Chapter 7 describes the limitations. Chapter 8 provides a short conclusion.

CHAPTER 2: KNOWLEDGE AS A PRODUCTION FACTOR

Innovations created on the Leiden Bio Science Park, such as bull Herman, are protected through patents. Patent law protects the newly developed innovations from copyright and provides a legal monopoly to produce and exploit these innovations for a maximum period of 20 years (Octrooicentrum.nl).

In the field of economics innovation is closely related to technology and knowledge. Simon (1973) stated that technology in its purest form is knowledge – knowledge to pursue our goals and solve our problems. The general perception is that innovations predominantly occur as a result of interactions between various actors, rather than springing from a single genius inventor's mind (Hakansson, 1987; Hippel, 1988; Lundvall, 1992 in Malmberg & Power 2005). The fact that biotechnology patents are predominantly registered as a cooperation between multiple inventors supports this perception. Austrian economist Joseph A. Schumpeter (1934) defined innovation as "neue Kombinationen" which means carrying out new combinations. According to the Schumpeterian perspective on innovation, interaction between inventors who possess dissimilar bodies of knowledge could lead to these new combinations which are described to be the ingredients for innovation. In the knowledge based biotechnology sector the focus is on the continuous radical and incremental innovation of products and processes.

2.1 Innovation and knowledge from a microeconomic perspective

This growing trend of continuous innovation can be further explored by taking a microeconomic perspective. The core of microeconomics is concerned with static efficiency by looking at the pricing system and the allocation of resources and rents accordingly. Allowing innovation to enter the equation changes this static model into a dynamic model, where the static efficiency is distorted and status quo changed (Koellinger, 2009) This leads to a model where the ability to respond quickly to a dynamic environment is increasingly valuable. Malmberg and Power (2005) place this development in context with previous models: *"This does not mean cost considerations are unimportant, but simply that the* *combined forces of market globalization and deepening divisions of labor make knowledge creation and innovation increasingly important."*

Krugman (1991) mentions that in neo-classical growth models, the long-run growth is determined by the exogenous factor of technological growth and that the concentration of knowledge leads to increasing returns and higher growth rates. This signals the importance of knowledge and indicates that firms need to seize opportunities that lie inside *and* outside the firm, as it is predicted to enhance their long-run growth rate.

To capture these outside opportunities the ability to learn and attract new knowledge has become an important or even unique sources of sustainable competitive advantage (Levinthal and March, 1993; Senge, 1990). Given these insights and industry developments biotechnology organizations don't leave their innovative ability to chance, but have routinized it in their activities. A study on drugs that entered clinical testing between 1990 and 2003 by DiMasi and Grabowski (2007) estimates that biotech firms spend an average of $615 million on research and development (R&D) per drug. Costs to perform clinical trials are estimated at an average of $626 million. The total costs to deliver a single drug to the market amounts to an estimated total of $1,241 million (DiMasi and Grabowski, 2007). These numbers suggest that in the biopharmaceutical industry the ability to generate new knowledge is a dominant factor in their production function.

2.2 The concept of knowledge further explored

Knowledge is a multilayered concept and is often used interchangeably with the term information. But the difference between the two is eminent; information can be perceived as a flow of messages, while knowledge is created by the flow of information in accordance with the commitment and beliefs of its holder (Nonaka, 1994). Alavi and Leidner (2001) elaborate on this by stating that knowledge is equal to authenticated information; *"It is personalized information related to interpretations, ideas, observations and judgments".* Dretke (1981) adds to this that *"the information one receives is always relative to what one already knows about the possibilities at the source".* These conditions regarding the authentication of information, implicates that knowledge is created by individuals (Nonaka, 1994). There are different types of knowledge residing in individuals, each requiring different means of learning and transfer. This can be illustrated with an example about learning how to ride a bicycle. You can acquire knowledge about how to handle the bicycle by reading an instruction manual or even watching an instruction video. But to actually be able to ride the bike, you need real life experience through practice and skills which cannot be learned by books or videos. Michael Polanyi (1966) signaled the same phenomena and put it this way: *"We know more than we can tell".* Clearly not all knowledge we possess can be transferred by explicit means. Following this insight, Polanyi differentiates the concept of knowledge into tacit and explicit or codified knowledge.

2.3 The concept of tacit knowledge

Tacit knowledge is rooted in action, commitment and involvement in a specific context (Nonaka, 1994). The example describing the process of learning how to ride a bicycle, illustrates the tacit knowledge elements residing in action and involvement. Compared to codified knowledge, tacit knowledge is more difficult to communicate and formalize. Therefore, to transfer this type of knowledge, face-to-face contacts and personal relationships are required (Breschi et al., 2005; Nonaka, 1994). The process of observation and imitation of specific craftsmanship routines can grasp the key of tacit knowledge. From this emerges that the concept of on-the-job training is for the most part concerned with the transfer of tacit knowledge (Audretsch & Feldman, 2003). Given these specific knowledge characteristics, tacit knowledge can be accessed without language and is often locally bounded (Nonaka, 1994).

2.4 The concept of codified knowledge

Codified knowledge, also referred to as explicit knowledge, is knowledge that is freely accessible and transmitted using formal, systematic language (Nonaka, 1994). This means codified knowledge can be expressed and recorded using words, numbers, codes, mathematical and scientific formulas and musical notations (http://www.businessdictionary online.com). Relevant codified knowledge for the biotechnology sector is published in scientific journals and patent databases.

Breschi et al. (2005) make an important note regarding the codification of highly complex and/or technical knowledge. They state that even though this type of often scientific and technical knowledge has been codified using formal language, a certain level of tacitness cannot be eliminated. Because the language used to codify the knowledge is so highly idiosyncratic, a certain vocabulary and level of experience with the subject is required to understand it correctly. Gaining access to this vocabulary and experience is limited to a small number of people, as the only way to learn is through prolonged studies and shared experiences (Oettl and Agrawal, 2008).

In this way, even fully disclosed knowledge can still not be used by actors residing outside the community, unless the actors inside the community decide to teach others how to use it (Hicks, 1995 in Breschi et al 2005). The limited group of people that is able to understand the codified information is called an *'epistemic community'* (Steinmueller, 2000). To enter the community and be able to understand the information transmitted, the possession of a certain level of tacit knowledge is required. Breschi and Lissoni (2009) argue that patents are an example of this type of codified knowledge: *"...patents represent a piece of codified information, but the knowledge stock they draw from is to a large extent tacit. In order to use that knowledge productively, one needs to have access to and interact with the individuals that have generated and still master it, that is, the patent inventors."* There is evidence that some firms purposely increase the tacitness of their knowledge to protect it from spilling over (Porter, 1998). The description of the LBSP biotechnology industry shows similarities with the epistemic community described by Steinmueller (2000). Also, the fact that knowledge is for a large part determining the organizations' competitive advantage, I expect them to deliberately increase the tacitness. Although patented biotech knowledge is freely accessible through online patent databases, I expect it to contain tacit knowledge elements. Therefore I assume that if organizations want to use patented biotech knowledge, they need to possess a certain amount of tacit knowledge. How organizations can gain access to this tacit knowledge is investigated in the next chapter.

CHAPTER 3: FROM ACCESSING TO UNDERSTANDING

In this chapter I describe the process that leads to formation of knowledge. Section 3.1 concisely describes the difference between information and knowledge. Section 3.2 describes the role of cognitive proximity in the process of individual knowledge creation.

Transforming information into knowledge

Transforming information into knowledge is a complex process and depends on various elements. When knowledge is transferred between two individuals, what the receiving actor receives at first is information. Only after the actor has authenticated this information it becomes knowledge again. I assume that the sending and receiving actor aspire for perfect mutual understanding, in order to avoid distortions in the knowledge creation process. However, in this transformation process differences in personal interpretation, ideas, observations and judgments, knowledge and beliefs can lead to different interpretations of a single piece of information (Alavi and Leidner, 2001). The severity of this distortion is determined by the extent to which the personal characteristics mentioned differ between sender and receiver. Therefore especially in an environment that deals with highly complex knowledge, such as the biotechnology sector, the relative importance of a shared tacit knowledge base increases. This tacit knowledge can often only be acquired through extensive study and experience (Oettl and Agrawal, 2008). Researchers in the field of economic geography have also studied this and other aspects of the knowledge transfer process. They evaluated the impact of a knowledge base distance between actors on the transfer of knowledge through the concept of cognitive proximity. The tacit knowledge or cognitive base varies for every individual on the planet.

3.1 Cognitive proximity

The concept of cognitive proximity is concerned with the knowledge base discrepancy between actors. An actors' acquired cognitive level is an accumulation of tacit and codified knowledge (Boschma, 2005). Cognitive proximity can therefore emerge when two actors share an educational background or have similar professional experiences. To facilitate an effective knowledge transfer a certain level of cognitive proximity between the two actors is necessary. The receiver needs to possess a certain absorptive capacity to identify, interpret and exploit the information received (Cohen and Levinthal, 1990). This absorptive capacity is determined by the knowledge base of the receiving actor. The knowledge base differences between the sending and the receiving actor represent their mutual cognitive distance. This knowledge or cognitive base consists of the technical and market competencies actors possess and have acquired while dealing with particular technologies and markets. If these competencies are not sufficiently shared, the costs for research and imitation of the shared knowledge will become too high (Boschma, 2005).

The receiving actor should posses a cognitive base that is close enough to the new knowledge in order to communicate, understand and process the information successfully (Boschma & Lambooy, 1999). However, the prevalence of either too much or too little proximity can cause the transfer of knowledge to become ineffective. Too little cognitive proximity means the absorptive capacity of the actor is not capable to understand the transferred knowledge. Too much proximity can also be harmful for the process of learning and innovation. Boschma (2005) summarizes three reasons why a certain level of cognitive distance should be maintained. First, the building of knowledge often requires dissimilar, complementary bodies of knowledge. Tapping into novel information may help arriving at new ideas and trigger creativity. Second, too much cognitive proximity may result in a cognitive lock-in. This can lead to a situation where firms only acquire new knowledge that is very close to their own cognitive base. When a firm ends up in the situation where they have difficulty unlearning habits or routines that have become successful in the past this is called the 'competency trap' (Levitt and March, 1996). Third reason has to do with knowledge spillovers. When the cognitive distance between two actors becomes very small they risk an involuntary spillover of knowledge. In a competitive situation this might not be wishful, as it may expose knowledge that gives away (some of) their competitive advantage.

Finding the optimal level of cognitive proximity is something each knowledge based firm should strive for when attempting to attract new knowledge. As the process of learning and knowledge creation is dynamic, the optimum will vary over time, per actor and type of knowledge involved. So when patented knowledge is transferred in the biotech industry I assume an optimal cognitive proximity between the two actors. On the other hand, a lack of knowledge transfer volume can indicate a disturbance in the knowledge transfer process. This can be the result of either a too large or too small mutual cognitive distance.

CHAPTER 4: THE TRANSFER OF KNOWLEDGE

In this chapter I describe the different aspects that play a role in the transfer of tacit knowledge. Section 4.1 describes the role of geographical distance. Section 4.2 explains the concept and impact of localized knowledge spillovers. Section 4.3 handles two knowledge transfer mechanisms. Section 4.4 describes the prevalence of local knowledge flows and their relationship to clusters. Section 4.5 describes the knowledge diffusion mechanism of labor mobility.

4.1 Geographical distance

As it is possible to store and publish codified knowledge in journals and databases, codified knowledge can be disclosed to anyone willing to search for it. For an effective transfer of pure codified knowledge, geographical proximity is therefore not a requirement. Or in other words, the transfer of codified knowledge is not affected by the geographical distance between the sender and receiver.

The opposite is true for tacit knowledge. Given the highly contextual nature and difficulty to codify tacit knowledge, it is best accessed through personal, repeated face-to-face contact (Audretsch, 1998). Establishing personal relationships and face-to-face contacts requires a certain level of geographical proximity (Breschi et al., 2005). Being geographical proximate thus enhances the local transfer of tacit knowledge. Geographical proximity may even facilitate an involuntary knowledge flow, as the knowledge transfer often occurs in an informal setting (Audretsch, 1998). The occurrence of uncontrolled knowledge spillovers is addressed through the concept of localized knowledge spillovers which is discussed in the next paragraph.

The transfer of tacit knowledge over a large geographical distance is more difficult to establish, as it becomes increasingly difficult to establish repeated face-to-face contacts and build the personal relationships described to facilitate a successful tacit knowledge transfer. For patented biotech knowledge with tacit elements this is an important issue. As the degree of tacitness determines the difficulty to transfer the knowledge over larger distances.

4.2 Localized knowledge spillovers

Adam B. Jaffe (1989) introduced the knowledge production function and accordingly the concept of localized knowledge spillovers. In his research he measured the effects of local R&D performances of universities on the number of private patent applications. He found evidence of a positive association between knowledge inputs and innovation outputs at the level of states, regions and cities. The evidence of these pure technical externalities leads to the introduction of the phenomena of localized knowledge spillovers (Breschi et al, 2005). The knowledge that spills over is highly contextual and difficult to codify, which are tacit knowledge characteristics (Breschi & Lissoni, 2001). Although these knowledge spillovers are not paid for by their recipient, they can nevertheless posses economic value in several different applications. Therefore these spillovers represent a public good (non-excludable & non-rival), but a local one given the transfer difficulties they possess (Audretsch and Feldman, 2003). These difficulties lie in gaining access to these localized knowledge spillovers as the spillovers are of a tacit nature. Audretsch and Feldman (2003) also state: *"the marginal cost of transmitting knowledge, especially tacit knowledge, is lowest with frequent social interaction, observation and communication."* So through frequent social interaction, observation and communication, tacit knowledge is transferred effectively. When actors share an increasingly larger base of tacit knowledge this increases understanding (or in other words decreases the cognitive distance) between these actors. So once personal relationships are established and face-to-face contacts become more intensive this may increase the local cognitive proximity as result of the geographical proximity. Following Boschma (2005) and Nonaka (1994) this increasing ability to understand the information available in the local environment decreases the number of distortions in the knowledge transfer. Therefore we argue that being geographical proximate might very well increase cognitive proximity which in turn reinforces the flow of tacit knowledge spillovers, in the local environment.

Furthermore, accessing local knowledge externalities reduces the costs of scientific discovery and commercialization. Traditionally many firms tend to let their location depend on the availability of resources. This suggests that innovation driven industries, where innovative activity and tacit knowledge play an important role, will show a higher tendency to spatially cluster (Feldman, 1994). The LBSP is a lighting example of an industry where innovation and tacit knowledge play an important role. The local presence of major research institutes such as the LUMC and other biotech organizations creates a snowball effect in the attraction of new organizations to the park.

4.3 Knowledge transfer mechanisms

There are many different ways to share the different types of knowledge. The focus in this thesis lies on codified knowledge with tacit knowledge elements, such as the biotech knowledge stored in patents. As codified knowledge flows freely to anyone willing to search for it, the tacit knowledge elements are more difficult to transfer. Moen (2005) concludes that tacit knowledge flows most effectively through two mechanisms, namely social networks and labor mobility.

To start with the first mechanism: a social network consists of an individual's total set of personal ties or relationships. These ties are established through personal interaction. Residing in a network can therefore foster the transfer of tacit knowledge through personal interaction with individuals in the network. In this thesis the flow of knowledge is analyzed from a dyadic perspective, the social network mechanism will therefore not be further discussed. However, the role of the social network in the diffusion of tacit knowledge is related to the formulation of hypothesis 1. If a local social network contains relevant tacit knowledge this can motivate an organization to co-locate inside this network. The phenomenon of co-locating has evolved over the past decades into the emergence of clusters as the LBSP. Porter (1994) already concluded that the economic map was dominated by clusters.

Moen's (2005) second mechanism concerned the transfer of tacit knowledge through labor mobility. Nonaka (1994) stated that knowledge resides for a large part inside the individual. Mobility of an individual implicitly means the knowledge that individual possesses travels along. Labor mobility as a knowledge transfer mechanism is therefore perceived to be an increasingly important knowledge resource (Breschi & Lissoni, 2009). Labor mobility of LBSP inventors facilitates them to establish face-to-face contacts and build personal relationships with their co-workers which are mentioned as conditions for an effective tacit knowledge transfer.

4.4 Local knowledge flows; locating inside a cluster

Reducing the geographical distance between competitors, industry related firms and research institutes is predicted to increase the potential access to localized tacit knowledge (Boschma, 2005). Organizations operating in high tech industries therefore often choose to locate in a cluster. Clustering or co-locating is the spatial aggregation of a certain type of economic activity (Porter, 1998). Over the past decades clusters have more and more been touted to be a major driver of economic development in the current knowledge based economic landscape (OECD, 1996). Generating new knowledge and the role it plays in attaining productivity growth has increased firms attention to circumstances and opportunities that reside outside the firm (Antonelli, 1998). By co-locating companies acknowledge that competitive advantages are not only to be attained inside the firm but more and more in the interaction with other firms (Barney, 1991). In the current perspective a lot of firms choose to locate inside a cluster as they expect it to be a factor that can enhance their productivity and innovative ability. Clustering is predicted to enhance inter- and intra firm contact as it allows the firm access to the local social network. At the same time labor market pooling effects can occur through the geographical aggregation of industry specific workers. This delivers benefits for both the worker and the firm, as it lowers the probability of unemployment and labor shortage through the creation of a regional specialized labor market (Marshall, 1920). A vibrant local labor market will attract specialized workers. These workers can improve the local knowledge base. The availability of a vibrant regional labor market for engineers, scientists and workers is therefore regarded as an important asset for the process of knowledge diffusion (Almeida and Kogut, 1999). Both the local social network and the increased possibilities for labor mobility are the two mechanism Moen (2005) reported to be effective tacit knowledge diffusion mechanisms.

I expect the combined effects of locating in a cluster such as the LBSP to play an important role in the knowledge diffusion. Given the presence of a local social network that facilitates local knowledge benefits I expect local developed knowledge also to diffuse local, where local means inside the cluster. This derivation leads to the first hypothesis:

Hypothesis 1:

H_0: Knowledge produced in the LBSP cluster is more likely to diffuse local

4.5 Labor mobility

Scientists in the R&D intensive biotechnology sector have access to valuable firm specific knowledge. Through working with this knowledge, scientists acquire a great deal of on-the-job training (Audretsch & Feldman, 2003). Investments made in R&D departments target the creation of new knowledge; however this knowledge is for a large part embodied in the workers. If the worker switches jobs, this accumulated human capital travels with him (Moen, 2005). Codified knowledge is often legally protected, through for instance patents or copyright law. However, when a worker switches jobs, the transfer of tacit knowledge is difficult to restrain (Oettl & Agrawal, 2008). At the new firm scientists can freely distribute their previous acquired tacit knowledge through the face-to-face contacts and personal relations at their new workplace. Evidence on the role of labor mobility as knowledge diffusion mechanism is found by Song, Almeida and Wu (2003). They find a positive relation between firms' ability to access technologically distant knowledge from other firms and the recruitment of engineers. Another study on inventor mobility by Rosenkopf and Almeida (2003) found similar results; firms do not only benefit from individual workers' knowledge, but they also seem to gain increased access to knowledge from the mobile inventor's prior firm (Oettl & Agrawal, 2008).

Important difference between the social network as tacit knowledge transfer mechanism and labor mobility is that labor mobility seems a more manageable knowledge diffusion tool. Placing a worker with a certain base of tacit knowledge inside a company is a very effective way of transferring specific knowledge. Compared to transferring knowledge through the social network a worker first needs to switch jobs before he/she can diffuse the knowledge, whereas knowledge diffusion through the social network only needs people be in the same network. Given the implicit higher costs of labor mobility I expect that this mechanism is more used to transfer tacit knowledge over relative larger geographical distances. Because transferring knowledge over a small distance is also easy to establish through the social network. I expect that as the distance between the sending and receiving actor increases, the social network mechanism becomes less effective and labor mobility can fill up this gap. This leads to the second hypothesis:

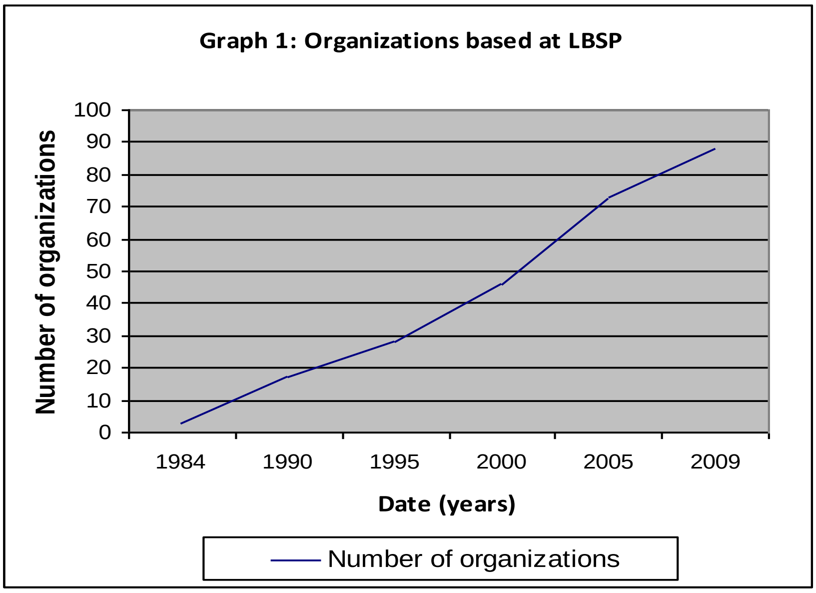
Hypothesis 2:

H_0: As the LBSP knowledge diffuses over a larger distance, it becomes more likely that this knowledge has been transferred through inventor mobility

CHAPTER 5: METHODOLOGY

The research part of this thesis set up as a case study. I investigate patented knowledge created by organizations residing on the Leiden Bio Science Park (LBSP). I track the diffusion of this knowledge and the role of labor mobility in the diffusion process. The knowledge diffusion is described in detail and I perform a simple test to find out whether this diffusion is locally biased. By comparing the knowledge diffusion path with the labor mobility of the inventors, I find out where the designated knowledge flows to and if this knowledge is transferred through the mobility of inventors. The focus of this chapter is on the localization question and knowledge diffusion preceded by inventor mobility.

This chapter is structured as follows; Section 5.1 concisely describes the Leiden Bio Science Park. Subsequently, section 5.2 continues with elaborating on the data sources and discusses the validity of these sources. Section 5.3 includes the construction of the dataset and provides a thorough description of the variables. section 5.4 deals with the statistical methods used in the analysis.



.

5.1 The Leiden Bio Science Park

The knowledge diffusion and labor mobility investigated in this thesis all stem from the LBSP. This bioscience cluster is located in the west of the Netherlands, in the hearth of the city of Leiden. The park was founded the 4th of April of 1984, after the Academisch Bedrijven

Centrum (ABC) was established on the park. As for April 2009 the park covered a total of 110 hectare, hosting 88 organizations and employing around 15.000 workers (http://www.leidenbiosciencepark.nl). Currently this biopharmaceutical cluster is dominated by medicine and life science organizations with expertise in the field of biotechnology for healthcare and/or biopharmaceutical technologies and products. Biotechnology uses (parts of) organisms in the development process.

Out of the 88 organizations located on the park there are two under foreign ownership; Genencor from Denmark and Centocor from the United States. Graph 1 show the remarkable growth of the park over the past 26 years. Currently it is the largest biotechnology cluster in the Netherlands and it's believed to be in Europe's top 5. The LBSP was rewarded the Menzis Award for Best Business Park 2009 in the Netherlands.

Public organizations play an important role on the park, such as for example the Leiden University Medical Center (LUMC). This is a leading institute in the international field of biomedical research. It plays an important role in attracting and training high skilled talents for the biotechnology sector. The medical center perceives its collaboration with Leiden University and the LBSP as an unique opportunity for medical innovation. The LUMC employs approximately 7.000 workers.

5.2 Data resources

The data in this thesis can be roughly divided in two parts; data describing the knowledge and its diffusion and data describing the labor mobility of LBSP inventors. The first dataset (i) describes the designated knowledge; a set of 30 major LBSP patents, and accordingly its knowledge diffusion path through the forward citations of these 30 major patents. The second dataset (ii) concerns the labor mobility of the inventors responsible for these 30 major LBSP patents

5.2.1 Knowledge and knowledge diffusion

The LBSP patent data has been extracted from the Organization for Economic Co-operation and Development's (OECD) REGPAT database. This database includes specific patent data that is linked to specific regions on the base of the addresses of patent applicants and inventors. The organization that centrally organizes patent registration is not the OECD but the European Patent Organization. This is an intergovernmental organization acting on the basis of the European Patent Convention since 1977. The European Patent Organization consists of two bodies, the European Patent Office (EPO) and the Administrative Council. The EPO intensively cooperates with the patent and trademark offices from the USA, Japan, Korea and China.

Forward citations are not included in the REGPAT database, but can be found in EPO's online patent database; Esp@cenet. In this database each patent has an individual page where all the patent's specifications are publicly available. This individual page can be traced through the patent's unique Publication Number, a code assigned to a patent by the EPO. The patent page provides, among other things, names and addresses information of the applicant(s), the technological classification of the patent, the inventors responsible for the invention and linkages to forward citing patents. By following these forward citation links in Esp@cenet it is possible to manually retrieve the forward citations' characteristics from their Esp@cenet page.

5.2.1.1 Data validity

The REGPAT and the Esp@cenet database both source their data from the European Patent Office. This intergovernmental institution is the executive arm of the European Patent Organization and is supervised by the Administrative Council. Given the fact that this is an independent organization where the executive and supervising bodies are separated, the validity of data provided by this institution is considered to be valid.

5.2.2 Labor mobility

Second part of the research is concerned with the job mobility of the inventors of the 30 major patents. Inventor names are extracted from the REGPAT database. To track the inventors' career I use web based resources. First online resource is the social media website LinkedIn. The users of this website are predominantly professionals, operating in all possible sectors. Individuals can place their résumé on their personal LinkedIn page and add colleagues, classmates and friends to their personal network. This allows others to track their careers and at the same time maintain their (professional) network. As for September 2010 this virtual network site is used by more than 80 million professionals over 200 countries to maintain and build their professional network (http://press.linkedin.com/faq).

In case inventors have not joined the LinkedIn community I consult two other online resources; first is the online patent database Esp@cenet and second is an online search through the Google search engine. The Esp@cenet database provides the option to search the database for inventor names. Performing this search results in an overview of all patents a specific inventor is registered on. Accordingly it provides the corresponding list of applicants he/she has been employed and has applied for a patent. Because it is not likely an inventor requests a patent at every organization they work for, so to increase the completeness of the data I perform an additional check through the search engine Goolge.com.

Search engines provide labor mobility information from many different sources. To ensure the credibility of the data I only use information distributed through professional organizations homepages and online biomedical and business journals.

5.2.2.1 Data validity

When performing web based research the validity of the collected data is difficult to verify as internet resources are easy to corrupt. By only using online information distributed through acknowledged organizations as LinkIn, Esp@cenet and websites of scientific journals and professional organizations I assume the provided information to be valid. An additional problem that the data collected from LinkedIn deals with is résumé fraud. Various research performed worldwide indicates that up to 50% (!) of all résumés contain discrepancies (www.intermediair.nl; www.cpai.com). Recent research performed in 2010 by the leading British pre-employment screening firm Powerchex revealed that 15% of a group of 5.858 British job applicants for the financial sector included discrepancies on their résumé (Powerchex annual pre-employment screening survey, 2010).

Main discrepancies concerned employment dates (33%), directorship (26%), academic record or qualifications (17%) and bankruptcy (12%). If I link these findings to my own data the discrepancies on employment dates could bias the data, as the exact date of job migration is important for this research. The other discrepancies are not likely to affect the data collected in this research as directorship, academic records and bankruptcy are no characteristics used in this research. Three remarks have to be made concerning the Powerchex research in comparison to the LBSP situation.

First of all, the majority of the LBSP inventors involved have obtained a master degree and often even a professorship at a university. According to the Powerchex survey a higher level of education is likely to decrease the likelihood an individual shows discrepancies on their résumé. Second remark is that the biotech inventor community is much smaller compared to the community operating in the financial sector on which the Powerchex research is based. A smaller network increases the risk of discovery of discrepancies. This will most likely increase the hesitation of employees to show discrepancies on their résumés. Last remark concerns the public availability of the LinkedIn résumés. These résumés are online available in contrast to the résumés used for the Powerchex research. This will most likely increase individuals' hesitation to add discrepancies as the risk of discovery increases. Based on these three counterarguments I expect that actually less than 15% of the résumés used in this research contain discrepancies and that the data collected online will be valid within reasonable boundaries.

The validity of the Esp@cenet database is not questioned as it provides patent data directly sourced from the European Patent Office.

Homepages of the inventors' (ex) employers and biomedical or business journals are also perceived to be reliable data resources. I assume that the organizations have no incentive to add discrepancies to their contents. Online irregularities are easily detected by third parties and this would cause damage to their reputation.

5.3 Construction of the dataset

This paragraph is divided in three parts. First part describes the construction of the variables describing the knowledge generated on the LBSP. Second part deals with the diffusion of this knowledge and final part entails the variables that describe the labor mobility of the inventors.

5.3.1 LBSP knowledge

The dataset includes knowledge protected by patent law, filed in the period between 1985 (01-01-1985) and the 1st of November 2007 (31-11-2007). Only patent applications from LBSP based organizations are considered. Applications filed before 1985 are left out, because patents filed in the first months of the park's existence predominantly are developed while the park did not existed yet and are therefore not produced on the park. The patents are distributed over the cohorts based upon their application date. The application date is the date the patent was first registered at the local EPO. Applying these selection criteria and removing missing patents from the REGPAT database results in a set of 496 LBSP patents (full list available on request). Because tracking the knowledge diffusion of 496 patents is beyond the scope of this thesis, therefore I narrow the focus of this research on the knowledge diffusion of the thirty most cited patents out of the total 496. I pick the most cited for three reasons:

(i) Several studies have indicated that the number of forward citations a patent receives can be used as a proxy for the technological importance, as well as social and economic value of a patent. (Almeida and Kogut, 1999; Albert et al. 1991, Carpenter et al. 1981). Where more citations indicate a higher value.

(ii) Constructing an overview of the most valuable knowledge created on the park and its diffusion is valuable information for organizations on the park and the regional/national public institutions supporting the LBSP. It can also add to the scientific literature for regional economists and economic geographers interested in cluster theory, knowledge development and diffusion.

(iii) In order to perform statistical analysis a sample of thirty observations is generally perceived as the minimum to get significant results. These thirty major patents will subsequently deliver a patent citation sample large enough to perform significant statistical analysis.

An important remark has to be made about operating counts research on forward citations. Older patents will have received relatively more citations than otherwise identical patents (Marco, 2006). This makes it difficult to exactly establish the value of the knowledge involved. However, the aim of the research is not to establishing the exact value of the knowledge involved, but merely to distinguish knowledge that is valued above a certain threshold and track its diffusion. Using these selection criteria implicates that I analyze the diffusion of knowledge with a certain technological, social and economic importance which narrows the scope of the research.

Applying the criteria mentioned above results in the following variables that capture characteristics of the pivotal LBSP knowledge:

<u>Variable 1; applicant name</u>

Applicant name(s) of the major patent. This information extracted from the REGPAT database.

<u>Variable 2; application date</u>

Application date of the major patent. This information is extracted from the REGPAT database.

<u>Variable 3; applicant address</u>

The address of the applicant organizations. The address of all applicants of the thirty major patents is Leiden. Also when it involved an application in cooperation with a firm from outside the LBSP I assume the main knowledge development occurred in Leiden. This information extracted from the REGPAT database.

Variable 4; applicant type

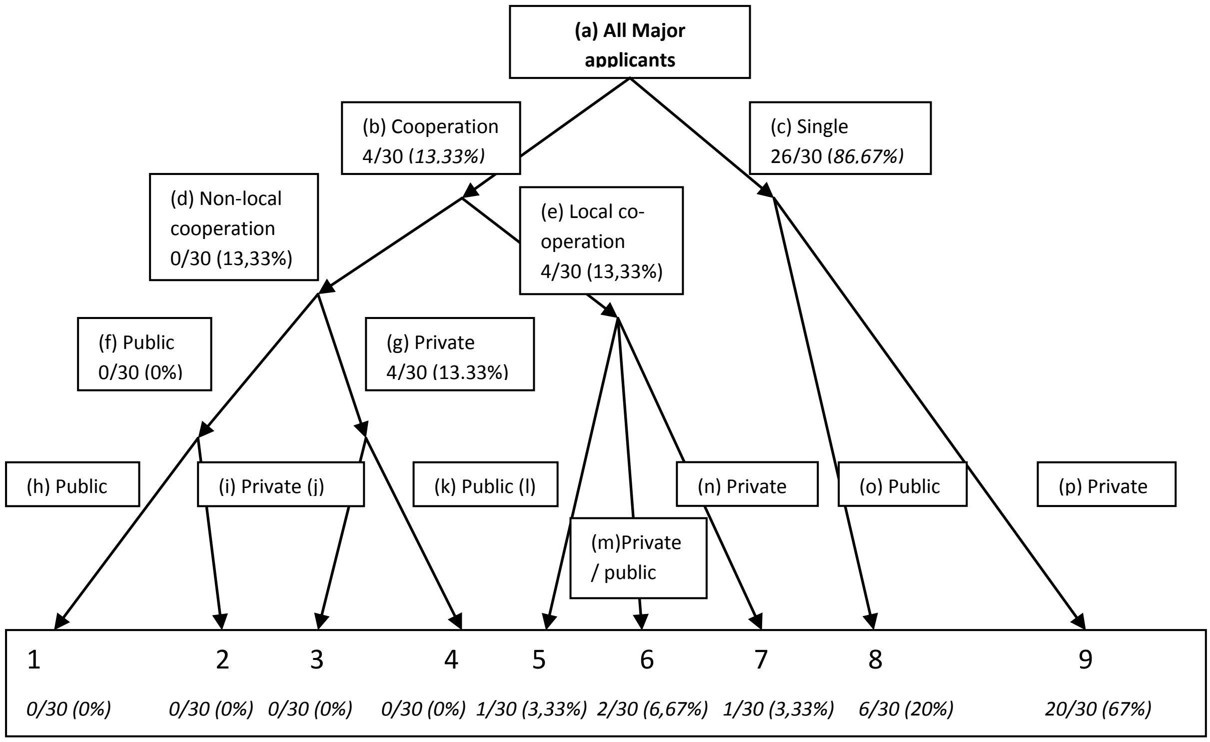
Dummy variable that captures the organizational form of the applicants behind the thirty major patents.

The applicants are differentiated on the base of the following characteristics:

- Single applicant or cooperation of applicants. Whether the patent was filed by a single applicant or the result of a cooperation of multiple applicants.
- Public or private organization. Universities, public research institutes, (academic) hospitals and other government funded organizations are labeled as public organizations. All others are labeled private.
- Organization residing on or outside the LBSP. Whether the organization was located on or outside the LBSP at the time of patent application. This differentiation only applies if the application involves a cooperative effort of at least two organizations.

This differentiation is visualized in diagram 1 and results in nine different types of applicants. The first box (a) represents the total population of applicants. Step one differentiates between applications filed as a (b) cooperation of two or more organizations and applications of a (c) single organization. For the patents filed by a single organization there is one more differentiation left. It concerns either a (p) (single) private organization; applicant group 9, or a (m) single public organization; applicant group 8. Next differentiation is between a (e) local and a (d) non-local cooperation. Local cooperations can be differentiated in three groups: (l) a collaboration between two local public institutions; applicant group 5. A local collaboration between a (m) private and a public organization; applicant group 6. And third, a cooperation between two private organizations; applicant group 7. If the cooperation concerned a (d) non-local collaboration, next differentiation deals with the question whether the LBSP organization involved in the local cooperation is a (f) public or (g) private organization. If it concerns a (f) public LBSP organization, next differentiation concerns the outside firm: either (h) public organization; group 1 or a (i) private organization; applicant group 2. If the LBSP organization is (g) private, the same differentiation applies concerning the outside LBSP organization in the cooperation: it's either (k) a public organization; applicant group 4, or a (j) private organization; applicant group 3.

Patent applicant figure 1: Patent applicant groups



5.3.2 Knowledge diffusion

Forward citations are registered by the EPO and eventually displayed in the Esp@cenet database. Data on the forward citations is manually collected from the Esp@cenet database in the period between the 10th of June 2010 and 24th of June 2010. By following the citation linkages in Esp@cenet I manually retrieved the characteristics for the knowledge diffusion dataset.

When handling the knowledge diffusion data it is important to realize that there are two ways a forward citation can be registered. First possibility is when the patent applicant acknowledges the forward citation in the application. Second possibility is that the patent bureau adds the citation. After a patent application is filed at the patent bureau, a special patent examiner performs a thorough check to find out whether the content of the patent has a relationship with a previous filed patent. If similarities are found, the examiner can autonomously add a citation to the application. This can lead to the event where an invention is developed without the scientist being aware of this previous developed, related knowledge, but still a forward citation is added to the patent application by the patent bureau. In this situation a forward citation is added, but an actual knowledge flow is absent.

So when using patent citations as a proxy for knowledge diffusion, one must realize that not all citations represent an actual knowledge flow. The scope of possible events, when performing research where forward citations are used as a proxy for knowledge flows, are categorized by Jaffe et al (1993). They arrange the possible events into the following three categories:

1. Knowledge flows accompanied by a citation.
2. Citations without an actual knowledge flow.
3. Knowledge flows without a citation.

The first category entails the events I aim to capture in the dataset. However, it cannot be excluded that category two events are also included in the data (Jaffe et al., 1993). There are two possible situations that can result into a category two event. First situation is when the patent bureau adds a citation. I expect this type of category 2 event to occur random. Therefore I expect the occurrence of this type of event to follow a normal distribution in the dataset. This will 'normalize' the data distribution and therefore level out deviating effects in the results. Based on this assumption I expect that effects found in the data in fact are stronger than they appear to be in the results.

Second possible event involves self-citations. Self-citations represent a situation where a patent applicant cites an own, previous filed patent. This leads to a citation without an actual knowledge flow, as the inventor already possessed the knowledge involved. Self-citations are registered in the major patents data and corrected for in the analysis.

Category three represents knowledge flows that do not leave a paper trail. As not all research output is patented, a large fraction of new developed knowledge remains traceless. This non-codified knowledge will not be represented in the dataset. However, if knowledge is not patented this does not automatically mean it is not valuable. (Koellinger, 2009)

To be able to track the knowledge diffusion I collected and calculated the following variables relating to the citing patents:

<u>Variable 5; name of citation applicant</u>

Applicant name of the forward citation. This information extracted from the Esp@cenet database.

<u>Variable 6; application date of citation</u>

Application date of the citing patent. This information is extracted from the Esp@cenet database.

Variable 7; time lag

This variable measures the time-lag between the application data of the major patent (application data) and application date of the forward citation (application date citation) in days. Leap years are not considered. This information is extracted from the Esp@cenet database.

Variable 8: country

Country of residence of the forward citing applicant. This information is extracted from the Esp@cenet database.

Variable 9; address

City of residence of the forward citing applicant's organization(s). This information is extracted from the Esp@cenet database.

Variable 10; distance

Variable that captures the geographic distance between the major patent applicant and the forward citing patent applicant. The distance is calculated by filling out the address and country of the major patent (which is always Leiden) and address of the citing patent on the website http://www.gpsvisualizer.com/calculators. This website provides a tool that calculates the geographic distance as a straight line between two locations.

Variable 11; local diffusion

This dummy variable captures whether the forward citing organization resides on, or outside the LBSP. Thereby it indicates whether it is involves a local or non-local knowledge flow. This variable is the dependent variable for hypothesis 1.

5.3.3 *Labor mobility*

I track the inventors' job mobility from the date their major LBSP patent is listed. This means their first employer is the applicant of the major patent they are listed on in the REGPAT dataset. Employment history prior to this stage is not taken into consideration. Job titles are not taken into consideration as I assume that regardless of the job description the inventors' knowledge will diffuse inside the new organization. Characteristics investigated are the timing of the labor mobility and the geographic location of the new organization. Timing of the job mobility is registered by looking at the order of employment. The starting organization is labeled 1 and the next organization 2 etc. If an inventor holds positions at different organizations simultaneously, the starting date of both employments will decide the order.

These results are recorded in a matrix where on the y-axis (vertical) the inventors are listed and on the x-axis (horizontal) the organizations.

Next step is to check whether the labor mobility complies with the knowledge diffusion. Now by comparing the inventor mobility matrix on similarities with the knowledge diffusion matrix it is possible to manually distinguish cases where the inventor and knowledge moved to the same organization. If these events occur simultaneously and the inventor moved to the organization before the citation was registered, I assume the mobile inventor was responsible for the knowledge diffusion.

For the dataset a dummy variable is created that reports whether similarities are found between the labor mobility of the inventor and the knowledge diffusion:

<u>Variable 12: labor mobility</u>
In case the knowledge flow has been preceded by the arrival of one of the inventors listed on the specific patent, the dummy variable reports a 1 if not a 0. This dummy variable is the dependent variable for hypothesis 2.

5.4 Method

For this research I formulated two hypotheses. First hypothesis investigates the knowledge diffusion of on the LBSP created knowledge and finds out to what extent this diffusion is locally biased. Second hypothesis investigates the role of labor mobility as a knowledge diffusion mechanism. For the first hypothesis the analysis is performed using simple probability calculation. The second hypothesis is tested using a probit regression.

5.4.1 Hypothesis 1

Dependent variable in this hypothesis is variable 11: local diffusion, which is a dummy variable. If the variable attains the value of one, this means the knowledge diffused local. In case it attains the value of 0, the knowledge diffused non-local. Given the scope of this thesis no control variable is constructed but a simple probability test is performed to check for localized knowledge diffusion.

The hypothesis stated:

Hypothesis 1: Knowledge produced in the LBSP cluster is more likely to diffuse local

The knowledge mentioned concerns the thirty most cited LBSP patents and local diffusion means the knowledge flows inside the LBSP cluster. I expect the likelihood of a LBSP patent to be cited local to be larger than the probability it is not local cited.

H_0: Prob (local diffusion = 1) > Prob (local diffusion = 0)

H_1: Prob (local diffusion = 1) ≤ Prob (local diffusion = 0)

So in order to establish the localization of the knowledge diffusion, I compute both probabilities by taking the LBSP sample as proxy for the LBSP knowledge diffusion and perform the test.

5.4.2 Hypothesis 2

Because the dependent variable 12 labor mobility is a binary or dichotomous variable a Cumulative Distribution Function (CDF) is used to execute the statistical analysis. Because I assume a normal distribution of the dependent variable, a probit analysis is preferred over a logit analysis.

The probit model is defined as: $\boldsymbol{Pr\ (y=1|x) = \Phi(xb)}$

Where y is the dependent (binominal) variable and x is the dependent variable regressed in the equation. The Φ stands for the standard cumulative normal probability distribution and xb is called the probit score or index. Because we assume a normal distribution for xb, the interpretation of the coefficients requires thinking following the Z (normal quantile) metric. Interpretation of the dependent variable b is as following; an increase in the dependent variable of one, means the probit score will increase by b standard deviations.
However, because the calculations are made in Stata, the interpretation of the results is more straightforward because Stata calculates the effective parameters itself.

<u>Assumptions of the model:</u>

- Relation between the independent variables and the dependent variable is linear
- Homogeneity of variance of the dependent variable over levels of the independent variable
- Homoskedastic residuals
- Absence of outliers

The model is constructed to fit the regression. It can estimate the relation of the parameters with the dependent variable and provide a correct evaluation of the results for hypothesis 2.

Hypothesis 2: *As the LBSP knowledge diffuses over a larger distance, it becomes more likely that this knowledge has been transferred through inventor mobility*

In other words, I assume a positive relationship between independent variable distance that measures the distance over which knowledge diffuses, and the dependent variable labor mobility, which indicates whether or not the knowledge transfer is preceded through inventor mobility. This results in the following model:

Probit (labor mobility = 1) = $\alpha + \beta_1$ (distance) + ε_i

The hypothesis assumes a positive value for $\beta_{1,}$ as this determines the coefficient and thereby whether the relationship between the dependent and independent variable is positive of negative. The α is the intercept and ε_i is the residual. The hypothesis can therefore be summarized into:

H_0: $\beta_1 > 0$

H_0: $\beta_1 \leq 0$

CHAPTER 6: RESULTS

Chapter 6 presents the results and discusses the outcomes of the methods described in the previous chapter. Section 6.1 presents descriptive statistics for the total group of patents, the major patents and the labor mobility of the LBSP inventors. It also describes the origins of the LBSP knowledge used in this research. Section 6.2 contains the knowledge diffusion analysis and discusses results hypothesis 1, regarding the localization issue of the knowledge distribution. Section 6.3 presents impact of the labor mobility as a knowledge transfer mechanism and handles the results for hypothesis 2.

6.1 Descriptive statistics

In this paragraph I will give a brief overview of a set of descriptive statistics for all the LBSP patent data. The dataset started off with a total of 496 LBSP patents, filed in the period starting from 1985 until November 2007. From this total a sample of thirty major patents is selected by looking at their number of forward citations.

6.1.1 Knowledge origin; applicant definition

Before I proceed with the exposition of the knowledge diffusion process, I first present the organizations responsible for the creation of the knowledge. I expect differences in the construction of knowledge to result in mutual differences in the succeeding knowledge diffusion process. There are approximately 104 different organizations behind the complete group of LBSP patent applications in the period from 1985 until November 2007. The thirty major patent applications are filed by fourteen different organizations. Table 1 presents the results for the major and all patent group. The all patent group has been added to the table as a benchmark for the results of the major patent group. The patent applications are differentiated through a set of three applicant characteristics; first criteria concerns if the application is filed by a single organization or a cooperation of multiple organizations, second criteria describes if the organizations involved are private or public organizations, third criteria indicates if organization involved is residing on or outside the LBSP (this last characteristic is only applicable if the application involves a cooperation). This set of criteria resulted in a set of nine different applicant groups, described under the header ‘applicant type’ in the table.

First result that stands out in table 1 is the large majority of 86,67% of major patents filed by a single organization. Out of these twenty-six single applicants, twenty applicants are private organizations and only six public organizations. Out of the total group of thirty patent applications twenty out of the thirty applications involve a single, private organization. However, only 13,33% of the major patents are filed as result of a cooperation.

Second result that stands out is the absence of non-local cooperation for the major patent applications; all of the cooperations signaled are local. This is an interesting result as it could be an indication of localization effects. Theory predicts that locating in a cluster enhances the access to the local social network, which in turn increases mutual cognitive proximity and could lead to local knowledge spillovers and enhances (local) effective cooperation (Porter, 1990; Barney, 1991). From the set of local cooperations there is a small majority of public – private cooperations. This could indicate that matching different types of organizations leads to higher valued knowledge. However the mutual differences and number of observations are too small to validate this statement. If we compare these results with the outcomes of the all patents group, a few remarks can be made.

Table 1: Overview patents applicants

		Major patents		All patents	
Number	Applicant type	Number of patents	Percentage	Number of events	Percentage
1	Non-local cooperation; LBSP public & non-local public	0	0%	18	3,63%
2	Non-local cooperation; LBSP public & non-local private	0	0%	34	6,85%
3	Non-local cooperation; LBSP private & non-local private	0	0%	10	2,02%
4	Non-local cooperation; LBSP private & non-local public	0	0%	8	1,61%
5	Local cooperation; public & public	1	3,33%	6	1,21%
6	Local cooperation; public & private	2	6,67%	13	2,62%
7	Local cooperation; private & private	1	3,33%	4	0,81%
8	LBSP public organization	6	20%	115	23,19%
9	LBSP private organization	20	66,67%	288	58,06%
	Total	30	100%	496	100%

From the all patents group 14,11% of the patents represent non-local cooperations, against 4,64% representing a local cooperation. The fact that the all patents group, representing a significant lower average patent value, is involved in non-local cooperation is striking. This could indicate that as the value of the invention rises, the propensity of non-local cooperations decreases and vice versa the propensity of local-cooperation increases. A possible explanation for this phenomena could be that organizations don't want their valuable knowledge to spill over to competitors (outside the cluster).

Related to this phenomena could very well be the relative high propensity of the major patent group involving in local cooperations compared to the all patents group. Local cooperation leads to higher valued inventions. Or maybe there is a reversed causality, that if the invention's promises to be of high value they choose to operate with local partners.

6.1.2 Major patents and forward citations distribution

Table 2 provides the distribution of the descriptive statistics over the five cohorts for the thirty major patents and their matching forward citations. Per cohort an overview is provided of the number of LBSP filed patents, forward citations, self-citations and the mean number of citations received per patent. The number of citations and the mean are both corrected for self-citations. All data is split out in five year cohorts to enhance the analysis of the distribution over time. For instance cohort 1985-1989 means the time period runs from 1st of January 1985 until the 31st of December 1989.

Table 2: Distribution of major patents and matching citations

Cohort	Patents	Matching citations	Self-citations	Citations minus self-citations	Mean	Mean (corrected for self-citations)
1985-1989	5	25	5	20	5	4
1990-1994	6	63	9	54	10,5	9
1995-1999	15	81	26	55	5,4	3,666667
2000-2004	4	20	4	16	5	4
2004-2007	0	0	0	0	0	0
Total	30	189	44	*145*	6,3	4,833333

Cohort 1995-1999 represents the highest number of patents, citations and mean with a total of fifteen patents, eighty-one citations and a mean of ten and a half citations per patent. These numbers suggest that this is the LBSP's most productive period so far in terms of knowledge production with social, economic and technological value. However, this cohort also holds the highest number of self-citations with twenty-six events, which bias the data. Self-citations appear in the data as a local knowledge flow. However, as they represent inventors citing their own, previous filed patent they don't represent an actual knowledge flow. After the correction for self-citations cohort 1995-1999 continues to represent the highest number of citations, but the mutual differences have decreased enormously. After the correction for self-citations cohort 1990-1994 now represents the highest mean, with an average of nine citations per patent. These outcomes are unexpected given the prediction of Marco (2006) who stated that older patents are likely to have obtained more citations than otherwise identical patents. If we look at the corrected mean, we see that knowledge from the second cohort shows the highest average, followed by knowledge from the first and fourth cohort, thereafter the third cohort and finally the fifth cohort. However, the fifth cohort has no observations and can therefore be left out of the analysis. The limited flow of knowledge for the first cohort can be explained with the fact that by then the park had not yet reached its full potential. As well the time-lag between having the idea and establishing a tangible invention takes a considerable amount of time. This development period in combination with the small number of organizations populating the park, can very well be the explanation behind the lack of new knowledge production in the first cohort.

These outcomes together with the limited flow of knowledge for the first cohort are unexpected given the prediction of Marco (2006), who stated that older patents are likely to have obtained more citations than otherwise identical patents. Possible explanations for this phenomena are the fact that by then the park had not yet reached its full potential (see graph 1). Also the time-lag between having an idea and establishing a tangible invention could play a role. The prevalence of a maturation period in combination with the small number of organizations populating the park, can very well be the explanation behind the lack of new knowledge production in the first cohort.

6.1.3 Labor mobility

The list of inventors considered consists of seventy-four individuals. Only eight out of these seventy-four inventors are female. Job mutations are considered starting from the organization they filed their major LBSP patent, registered in the REGPAT database. The group of inventors was involved in seventy-eight job mutations in the relevant period from 1985 until November 2007. This means that including their LBSP employment they filled a total of hundred and fifty-two different positions. The average number of switches for the inventors involved amounts to 1,05 switches per inventor or 2,05 different employments per inventor. Altogether the inventors worked at fifty-nine different organizations. The detailed inventor mobility matrix is available upon request.

Table 2 presents an overview of the 74 inventors' mobility. Again, the number of switches is measured starting from the position they held when they filed their LBSP patent. This means that if an inventor was involved in 8 switches, in fact he/she was employed at nine organizations. When organizations merged or taken over, this is not regarded as a switch of employment. If an inventor starts working part-time at a different organization but also remains employed in his current job, this is regarded as a switch. The relative distribution presented in table 2 is calculated by dividing the number of inventors that undertook a certain number of switches by the total number of inventors, times 100%.

Table 2: Inventor mobility overview

Number of switches	Number of inventors	Relative distribution
0	38	51,51 %
1	21	28,37 %
2	5	6,76 %
3	4	5,41 %
4	0	0 %
5	3	4,05 %
6	2	2,70 %
7	0	0 %
8	1	1,35 %
Total	74	100%

The majority of inventors did not switch jobs from the point their LBSP patent was filed. Second largest group in the distribution represents the inventors that switched job once. As the number of switches increase, the number of inventors involved decrease. With twenty-six out of the thirty major patents being filed before the start of 2000, this could be an indication that the level of job mobility is not too high for inventors operating in the biotechnology sector.

6.1.4 Time-lag & distance

Table 3 summarizes the distance and time-lag characteristics for the different cohorts. Total geographic distance traveled by the LBSP knowledge accumulates to 583260,24 km; a distance that equals circling earth more than fourteen times. The average geographical distance between the major patents and forward citations is 3766,82 km. The average time-lag between the filing date of the major patent and the filing date of the forward citing patent equals 1833 days, or five years, one month and five days (without taking leap-years into consideration). The average time-lag seems to diminish over time, this could be an indication that LBSP knowledge is diffusing on an increasingly higher rate.

Table 3: Time-lag and distance between patent - citation

	Cohort 1985-1989	Cohort 1990-1994	Cohort 1995-1999	Cohort 2000-2004	Cohort 2005-2008	All cohort average
Total distance to citation (km)	80569,94	178860,7	282621,8	41407,8	na	145815,06
Average distance to citation (km)	4.028,497	3.312,236	5.138,578	2.587,988	na	3766,82
Average time-lag to citation (days)	2.509,95	1.996,741	1.575,764	1.320,625	na	1850,77

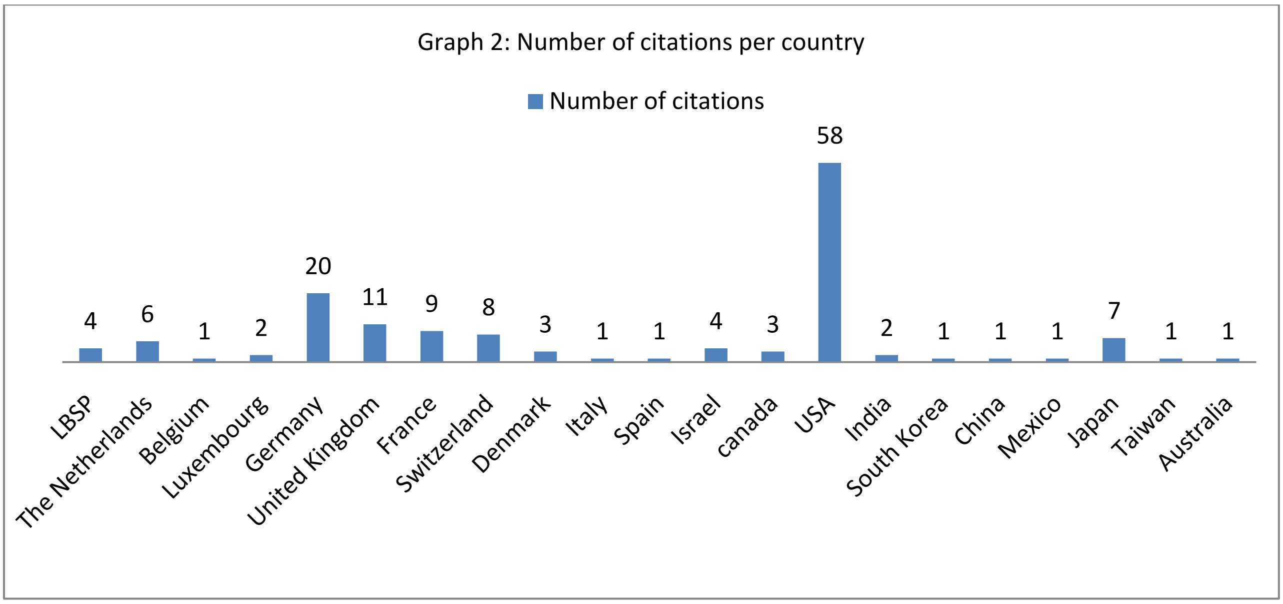
6.2 Knowledge diffusion

In this paragraph I describe the knowledge diffusion of the thirty major patents and present the results of hypothesis one. The diffusion is tracked using patent citations, as described in detail in the methodology. Self-citations are removed from the data and not considered in the analysis. The knowledge diffusion is thoroughly described per cohort, to be better able to signal trends and mutual differences. First, an overall impression of the geographical diffusion of the knowledge is provided. Subsequently these results will be discussed per cohort. The dataset consists of a total of 145 forward citations, excluding self-citations. These citations represent a knowledge flow from the LBSP (Leiden, The Netherlands) to the geographic location of the organization behind the forward citing patent. Only citations registered in the period 1st of January 1985 until the 1st of November of 2007 are considered.

6.2.1 Knowledge diffusion per country

The knowledge flows to twenty different countries, distributed over three continents. Graph 2 shows the distribution per country. The countries are sorted on the x-axis by their geographical distance to the LBSP. The y-axis represents the number of citations. LBSP citations represent a local knowledge flow; i.e. knowledge diffused to other organizations residing on the LBSP. The utter right side represents knowledge flows to Australia based organizations; the average measured distance to Australia amounts to 16.655 km.

When looking at graph 2 a few results stand out. The top ten citing countries account for 89,67 % of the total citations and that the top five accounts for 73,10 % of the total. Out of the top five citing countries only the USA is not located in Western Europe. A large majority of the knowledge flows to organizations residing in the USA, with at total of 40% of the citations. There are only four citations registered by other LBSP firms which correspond to 3% of the distribution.

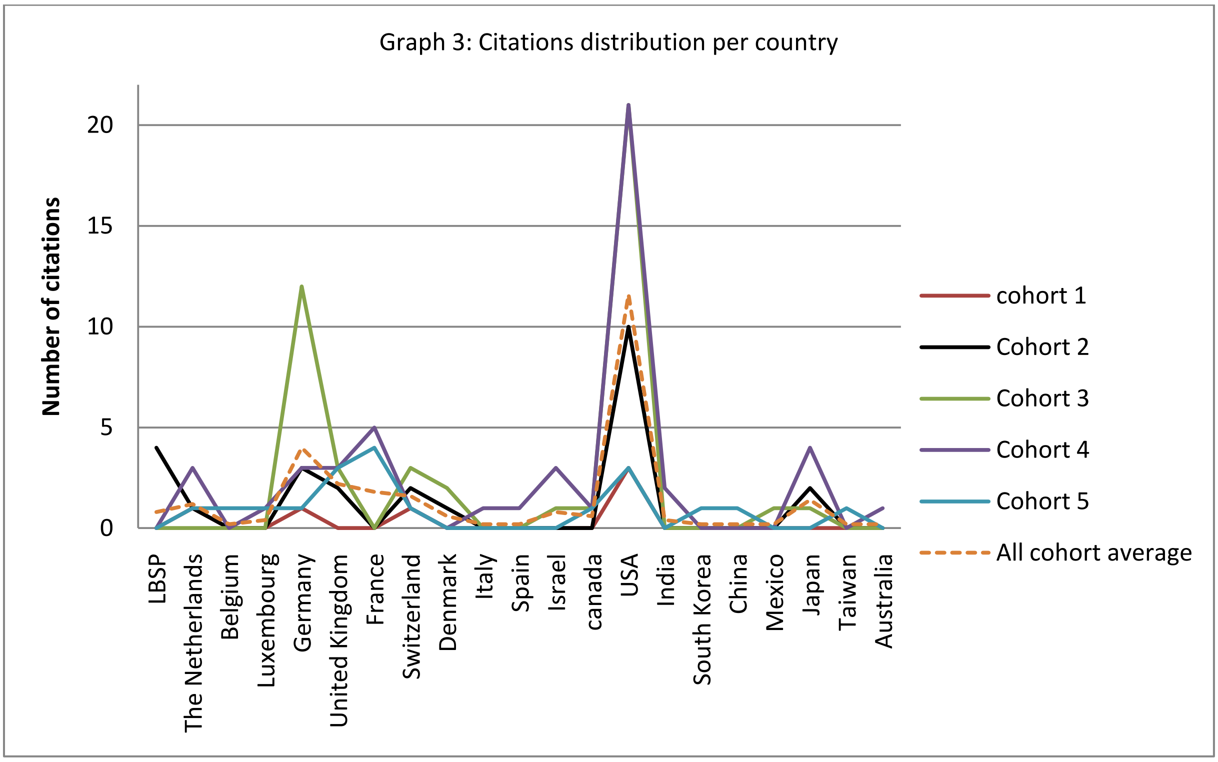


6.2.2 Knowledge diffusion per country, per cohort

The distribution of citations over the five cohorts is illustrated in graph 3. The countries labeled on the x-axis are again distributed according to the average distance to the LBSP. On the utter left side we have local citations and on the utter right side the Australian citations. For the overview of average distances see the appendix. Each of the five cohorts are represented in the graph, as well as the all cohort average.

The all cohort average per country shows that the highest number of citations are received from USA based organizations. Western Europe marks place two till five regarding the number of citations received. The difference between the USA with 58 citations and number two Germany with 20 citations is remarkably large.

Forward citations applied for in cohort 1 (1985-1989) are represented by the red line. This cohort entails only seven citations, which distribution don't substantially deviates from the all cohort average with citations a majority of citations by Western European and USA based organizations.



Forward citations applied for in cohort 2 (1990-1994) are represented by the black line. Cohort 1990-1994 deviates from the average distribution with a minor increase in the number of local citations. Rest of the knowledge distribution generally follows the all cohort average.

Forward citations applied for in cohort 3 (1995-1999) are represented by the green line. The cohort 1995-1999 shows a regional bias in its citation distribution, with a large peak for citations registered in Germany. Thereafter it has its largest peak for citations from the USA. with a total of twenty-one citations.

Forward citations applied for in cohort 4 (2000-2004) are represented by the purple line. This cohort entails the relative highest number of observations and that might be an important reason why its distribution for a major part complies with that of the all cohort average. It has its largest peak for citations from the USA and second largest for France. It is remarkable that the cohort has no local citation registrations, but a small peak for citations from the organizations in the Netherlands but outside the LBSP. What subsequently catches the attention is that the fourth cohort has small citation peaks for France, Israel and Japan.

Forward citations applied for in cohort 5 (2005- 2007) are represented by the blue line. The citation distribution of this last cohort shows a minor regional bias, with a peak for the Western European countries. USA citations are relatively underrepresented in this cohort.

6.2.3 Hypothesis 1

Testing overall localization of knowledge is executed through a simple odds distribution. The distribution and probability of local versus non-local citations is as following:

Table 4: Local knowledge diffusion (hypothesis 1)

	Local citations		Non-local citations	
Cohort	Citations	Probability	Citations	Probability
1985-1989	0	0	6	1
1990-1994	4	0,16	21	0,84
1995-1999	0	0	45	1
2000-2004	0	0	50	1
2005-2008	0	0	19	1
Total	4	0,0276	141	0,0724

Hypothesis 1: *Knowledge produced in the LBSP cluster is more likely to diffuse local*

H_0: Prob (local diffusion = 1) > Prob (local diffusion = 0)

H_1: Prob (local diffusion = 1) ≤ Prob (local diffusion = 0)

We find here that Prob (local diffusion =1) < Prob (local diffusion = 0) for all periods and also over the entire period. H_0 can therefore be rejected based upon these results.

From these results we can conclude that the considered knowledge from the LBSP is not locally bounded. In only four cases out of the 145 forward citations the knowledge flowed local. Surprising is that all local citations occurred in the cohort 1990-1994. The first hypothesis can therefore be rejected for all periods. In graph 1 we find a strong bias of knowledge flowing to the USA.

6.3 Labor mobility as knowledge diffusion mechanism

There are six cases in which the citation of a patent has been preceded by the mobility of an inventor. Table 5 shows the distribution of events, together with the calculated probability based upon the sample of thirty major LBSP patents. With a propensity of 0,0414 of the 145 citations the emergence of labor mobility as knowledge carrier seems to be not very eminent. The six events are equally distributed among local and non-local knowledge diffusions. This is a remarkable finding given the overall distribution of the citations. In the descriptive statistics we found that the majority of the knowledge flowed to the USA and that only a minor part of the knowledge diffused locally. Based on these findings I would expect the distribution of local and non-local labor mobility together with knowledge diffusion to follow this trend. However the equal distribution of local and non-local knowledge diffusion preceded by labor mobility suggests that labor mobility as a knowledge transfer mechanism occurs more often locally. This is the opposite of what I hypothesized.

Table 5: labor mobility & knowledge diffusion

	Labor mobility = 1		Labor mobility = 0		Total	
	Events	Probability	Events	Probability	Events	Probability
Local diffusion = 1	3	0,0207	1	0,0069	4	0,0276
Local diffusion = 0	3	0,0207	138	0,9517	141	0,9724
Total	6	0,0414	139	0,9586	145	1

Before we proceed with hypothesis 2 we first take a closer look at the six matching cases represented in table 6.

What is striking is that the six events are caused by only three scientists. Event 1 represent a knowledge transfer from a private LBSP organization to a private organization outside the LBSP, but still in the Netherlands. Event 2 also represents a transfer from a private organization but to a public organization in the state of Indiana, USA. Event 3 till 5 represent

Table 6: Overview of (Local diffusion = 1) occurrences

Matching event	Citing patent number	Filing date original patent	Scientist involved	Name original applicant	Citing date	Name citing applicant
1	14	13-9-1990	Elzen van den, Peter J. M.	MoGen	23-2-1994	Unilever (NL)
2	18	13-9-1990	Elzen van den, Peter J. M.	MoGen	9-9-1994	Unilever (NL)
3	51	7-6-1990	Woloshuk, Charles Peter	Syngenta Mogen	27-9-2004	Purdue University (USA)
4	132	29-10-1985	Hoekema, Andreas, Drs.	LUMC	23-3-1990	MoGen (LBSP)
5	133	29-10-1985	Hoekema, Andreas, Drs.	LUMC	21-9-1990	MoGen (LBSP)
6	134	29-10-1985	Hoekema, Andreas, Drs.	LUMC	23-3-1990	MoGen (LBSP)

a local knowledge flow. From the LUMC to the private organization MoGen. Event 3 and 4 have been cited on the same day but are independent patents. Interesting is that the three original patents stem from the first 6 years of the parks' existence.

The average time-lag between original and citing date is twenty-four days, eleven months and six years. The average time-lag of all LBSP patents was thirty days, eight months and four years. A longer time-lag could be regarded as an indicator for the necessity of labor mobility function as a knowledge transfer mechanism. But given the small number of observations this is not a strong indicator.

6.3.1 Hypothesis 2

The results of the probit analysis of the model are presented in table 7.

Table 7: Probit results hypothesis 2

Number of observations: 145

Labor mobility = 1	Coëfficiënt	Standard error	Z	P>|z|	[95% Confidence Interval]	
Distance	-0.001947	0.000637	-3.06	0.002	-0.0031962	-0.0006985

** Removing the residual from the regression increased the significance of the regression, therefore this value is not displayed in the table*

Probit (labor mobility = 1) = $\alpha + \beta_1$ (distance) + ε_i

The relation of the independent and dependent variable is negative, indicating that an increase in distance travelled by the knowledge decreases the likelihood this knowledge flow is preceded by inventor mobility. The value of the coefficient indicates that if the knowledge travels one extra kilometer the likelihood decreases with 0,001947. The z-value indicates the significance of the regression. The low P>|z| value represents the likelihood the value of the parameter is equal to 0 and is really low. Therefore the significance of the regression is reasonable high. However, important note has to be made concerning the sample size and the low number of events labor mobility = 1.

Hypothesis 2: A*s the LBSP knowledge diffuses over a larger distance, it becomes more likely that this knowledge has been transferred through inventor mobility*

The expected relation between distance and labor mobility = 1 to result in a positive parameter β_1. This could be translated into:

H_0: $\beta_1 > 0$

H_1: $\beta_1 \leq 0$

However, table 7 presents a negative value of -0,001947. Therefore the hypothesis can be rejected as an increase in distance decreases the likelihood of labor mobility as knowledge diffusion mechanism.

CONCLUSION

The research question of this thesis stated: “Where does LBSP generated knowledge flow to and to what extent does labor mobility play a role in diffusing this knowledge? Theory suggested that the knowledge would diffuse locally (Jaffe et al., 2003) and that labor mobility would play an increasing important role in the diffusion of knowledge (Song, Almeida and Wu, 2003). Based upon the conditions for knowledge transfer I hypothesized that as the geographic distance travelled by the knowledge increased, the likelihood labor mobility would function as the diffusion mechanism would increase. However, based upon the research conducted in this thesis, both hypotheses did not hold. The knowledge generated in the LBSP flows doesn’t flow locally but predominantly to the United States and to the Western European countries Germany and France. This result could possibly be explained by the statistic method used and the size of the sample.

The role of labor mobility as a facilitator of the LBSP knowledge diffusion is in this dataset quite modest. It facilitated knowledge diffusion in only 4% of the cases. Interesting finding is that however the impact of labor mobility was small, it had a relative larger impact for mobility within the cluster. This could indicate that the LBSP has a functioning local labor market of highly skilled inventors. Now back to the second hypothesis. Result of the second hypothesis indicated that when knowledge diffused over longer distances, the likelihood it has been preceded by labor mobility diminished. The effect I found indicated that for every extra kilometer the knowledge transferred the likelihood it was diffused through labor mobility decreased with -0,001947. This is an interesting finding as one would expect that the social network would function as diffusion mechanism for the smaller distances and labor mobility for the longer ones. However, other variables seem to play a role in this matter. It could be that the inventors are more reluctant to move over larger distances. This could be substantiated by the relative low level of job mobility of the LBSP inventors.

Last scientific attribution to this field of research lies in the data collection; the data collected on the inventors job mobility is more complete than previous conducted research (Song, Almeida, Wu, 2003; Rosenkopf and Almeida, 2003; Oettl and Agrawal, 2008). The availability of LinkedIn as an online job mobility database can prove to be an important research tool for future research. Around 50% of the inventors was registered at the LinkedIn website.

Other interesting result is the role of private organizations on the LBSP. More than 73% of the high valued knowledge diffusion descends from private LBSP organizations. This overrepresentation could be an indication that the cluster's innovative strength lies for a large part in the hands of private organizations. This knowledge could prove to be valuable for the organizations supporting and located on the LBSP and other high-tech clusters in Western countries.

LIMITATIONS

The conducted research in this thesis is subject to a number of limitations. Most limitations concern assumptions made and the research part of this thesis. In the second chapter of this thesis I follow the assumption of Breschi and Lissoni (2009) that patented knowledge contains tacit knowledge elements. Based upon this assumption I further derive that the biotech knowledge investigated cannot be transferred without gaining access to the inventors of this particular knowledge. This is a simplification that sets false expectations regarding the knowledge diffusion process. By disregarding further knowledge specifications such as technology class (IPC code), the size of the applicants organization and distribution of other expert organizations and clusters worldwide diminishes the significance of this research setup. Because the biotechnology sector is highly specialized and operates under high risk abnormal circumstances that need to be accounted for in designing the research and interpreting the results. Another limitation concerns the selection of the thirty major patents. By selecting the sample on the basis of forward citations the randomness of the sample is biased.

Second limitation lies in taking patent citations as a proxy for knowledge diffusion. The risk remains that citations without an actual knowledge flow are recorded in the data, which can bias the results. Also, the citations do not cover the full set of knowledge spillovers. By not tracking the labor mobility of inventors prior to their LBSP patent could mean an important set of data is missing. As Oettl and Agrawal (2008) mentioned; 'new firms can also tap into work experience obtained prior to the current job'.

The method used in hypothesis 1 is not very sophisticated. The lack of a control variable decreases the significance of the obtained results. The method used by Jaffe et al (1993) makes it possible to conduct thorough investigation of the localization knowledge flow, however this was outside the scope of this research given the already large amount of manual data collection performed. The same argument goes for hypothesis 2, again the lack of a control variable decreases the value of the results. However, the suggestions made for control variables by the literature were not feasible for this thesis.

REFERENCES

Abramovsky, L., Harrison R., and Simpson, H. 2007. University Research and the location of Business R&D. *The Economic Journal,* No. 117, pp. 114-141

Albert, M. B., D. Avery, F. Narin, P. McAllister. 1991. Direct validation of citation counts as indicators of industrially important patents. *Research Policy*. No. 20, pp 251–259

Audretsch, D. B., Feldman, M. P. 1996. R&D spillovers and the geography of innovation and production. *American Economic Review*. No. 86, pp 630-640

Audretsch, D. B. 1998. Agglomeration and the location of innovative activity. *Oxford Review of Economic Policy.* Vol. 14, No. 2. pp 18-29

Audretsch, D. B., and Feldman, M. P. 2003. Knowledge Spillovers and the Geography of Innovation. *The Handbook of Urban and Regional Economics.* No. 4

Alavi, M. Leidner, DE., 2001. Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues. *MIS Quarterly*, 25 (1): 107-146

Antonelli, C. 1998. Localized Technological Change, new information technology and the knowledge-based economy: The European Evidence. *Journal of Evolutionary Economics.* Vol 8, pp 177-198

Barney, J. 1991. Firm Resources and Sustained Competitive Advantage. *Journal of Management*, Vol. 17, No. 1, pp 99-120

Breschi, S. and Lissoni, F. 2009. Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography*, Vol. 9, pp. 439–468

Breschi, S., Lissoni, F. and Montobbio, F. 2005. The geography of knowledge spillovers: Conceptual issues and measurement problems. In Breschi, S. and Malerba, F. 2005. *Clusters, Networks and Innovation. Oxford,* New York: Oxford University Press Inc

Carpenter, M., Narin, F., Woolf, P. 1981. Citation rates to technologically important patents. *World Patent Information*. Vol. 3, No. 4, pp 160–163

Cohen, W. M. and Levinthal, D. A. 1990. Absorptive capacity: a new perspective on learning an innovation. *Administrative Science Quartely.* Vol. 35, pp. 128-152

DiMasi, J.A. and Grabowski, H.G. 2007. Managerial and Decision Economics. Vol. 28, pp 469–479

Dretske, F. (1981), *Knowledge and the Flow of Information*, Cambridge, MA: MIT Press

Feldman, M. P. 1994. *The Geography of Innovation.* Boston: Kluwer Academic Publishers

Grant, R. M. 1996. Prospering in Dynamically-Competitive Environments: Organizational Capability as Knowledge Integration. *Organizational Science*, 7, pp. 375-387

Hakansson, H. 1987. *Private Technological Behaviour: Co-operation and Networks*. London: Routledge

Hippel, E. V. 1988. *The Sources of Innovation*. Oxford: Oxford University Press

Jaffe, A. B. 1989. Real effects of academic research. *American Economic Review*. Vol. 79, No. 5, pp. 957-970

Hofer, C. and Schendel, D. 1978. *Strategy Formulation: Analytical Concepts.* St Paul: West Publishing Co.

Levinthal, D. A. and March, J. G. 1993. The Myopia of Learning. *Strategic Management Journal,* Vol. 14, pp. 95-112

Krugman, P. R. 1991. *Geography and Trade.* Cambridge: MIT Press

Lundvall. B. A. 1992. *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning.* Londen: Printer

Malmberg, A. and Maskell, P. 2002. The Elusive concept of Localization Economies: Towards a Knowledge-Based theory of Spatial Clustering. *Environment and Planning*. Vol. 34, pp. 429-449

Malmberg, A. and Power, D. 2005. (How) Do (Firms in) Clusters Create Knowledge? *Industry and Innovation*. Vol. 12, No. 4, pp. 409-431

OECD. 1996. The Knowledge-Based Economy. *Science, Technology and Industry Outlook*. Pp. 229-256, Paris

Oettl, A. and Agrawal, A. 2008. International labor mobility and knowledge flow externalities. *Journal of International Business studies*. No. 39, pp. 1242-1260

Sorensen, O. 2005. Social Networks and the Persistence of Clusters: Evidence from the Computer Workstation Industry. *In Clusters, Networks, and Innovation.* Breschi & Malerba, 2005. Oxford University Press, Oxford New York

Marco, A. C., 2006. The dynamics of patent citations. *Economic Letters*. Vol. 94, pp. 290-296

Marshall, Alfred. 1920. *Principles of Economics,* 8th ed. London: McMillan

Nonaka, I. 1994. A Dynamic Theory of Organizational Knowledge Creation. *Organizational Science,* Vol. 5, No. 1, pp. 14-37

Nonaka, I. Toyama, R., Konno, N., 2000. SECI, BA and Leadership: a Unified Model of Dynamic Knowledge Creation. *Long Range Planning,* Vol. 33, pp. 5-34

Polanyi, M. 1966. *The Tacit Dimension.* Londen: Routledge & Kegan Paul

Porter, M. 1985. *Competitive advantage.* New York. Free Press

Porter. M. 1990. *The Competitive Advantage of Nations.* New York: Free Press

Porter, M. (1998). Clusters and competition: New agendas for companies, governments, and institutions. In Porter, M.,. 1979. *On competition.* pp. 197-287. Boston: Harvard Business School Press

Schumpeter, J. 1934. *Capitalism, socialism, and democracy.* New York: Harper & Row

Senge, P. 1990. *The Fifth Discipline: The Art and Practice of the Learning Organization*. New York: Doubleday

Simon, H. A. 1973. The Structure of Ill Structured Problems. *Artificial Intelligence,* Vol. 4, No. 3-4, pp. 181-201

Steinmueller, E. 2000. Does Information and Communication Technology Facilitate ‘codification’ of knowledge?. *Industrial and Private Change,* pp. 361-376

Thompson, P., Fox-Kean, M. 2005. Patent Citations and the Geography of Knowledge Spillovers: A Reassessment. *The American Economic Review*. Vol. 95. No. 1, pp. 450-460

APPENDIX

Table 8: Full list of thirty major patents

Application id	Application number	Application date	Number of Citations	Applicant name 1	Applicant name 2
17296278	91202355	1990	21	GIST-BROCADES N.V.	Mogen Int
17294439	91200166	1990	16	BROCADES PHARMA B.V.	
17667643	97202523	1997	14	Universiteit Leiden	
17295432	91201344	1990	13	Syngenta Mogen B.V.	
17388920	92923499	1991	12	Syngenta Mogen B.V.	
17177113	89200736	1988	11	MOGEN INTERNATIONAL N.V.	
17528463	95202213	1994	11	Crucell Holland B.V.	
17665871	97200022	1996	11	Syngenta Mogen B.V.	
17667929	97202909	1997	10	ACADEMISCH ZIEKENHUIS LEIDEN	
17836541	99201278	1998	6	Crucell Holland B.V.	
17749238	98204482	1998	6	Introgene B.V.	
17838500	99203983	1999	6	Introgene B.V.	
15953964	1200321	2001	6	Cyto-Barr B.V.	
17033584	86201878	1985	5	Universiteit Leiden	
17596761	96203234	1996	5	OctoPlus B.V.	
17836776	99201593	1998	4	Crucell Holland B.V.	
17838212	99203578	1999	4	Universiteit Leiden	
15846721	203030	2000	4	OctoPlus B.V.	
15844848	200242	2000	4	Universiteit Leiden	Seed Capital B.V.
17075932	87200348	1986	3	Universiteit Leiden	
17125019	88201871	1987	3	NIJSSEN LIGHT DIVISION B.V.	
17179114	89202883	1988	3	H.B.T. HOLLAND BIOTECHNOLOGY B.V.	
17470053	94203630	1994	3	Fokker Space B.V.	
17527773	95201210	1995	3	Universiteit Leiden	
17668768	97204098	1997	3	Fokker Space B.V.	
17838419	99203878	1998	3	Crucell Holland B.V.	
17747225	98201693	1998	3	Introgene B.V.	
17837227	99202234	1999	3	Introgene B.V.	
15954222	1200711	2000	3	Crucell Holland B.V.	
17527629	95201003	1995	2	Fokker Space B.V.	

APPENDIX

Full list of all organizations the inventors have held employment

Table 9: Full list of organizations inventors held employement

Number	Organization name
1	Add2xBio
2	Octoplus
3	Antabio SAS, Toulouse France
4	Oxford University
5	Astrazeneca (werd Syngenta Mogen), Zoetermeer
6	Pharming
7	Audion Therapeutics
8	PlantZyme (joint venture of MOGEN and DSM), Enkhuizen
9	Batavia Bioservices BV
10	Polyvation
11	BioConsilium SARL, France, Toulouse
12	Profibrix
13	Brocades Pharma B.V.
14	Prosensa
15	CatchMabs, Wageningen
16	Proteonic
17	Crossbeta Biosciences
18	Purdue University, West Lafayette, Indiana, US
19	Crucell (patents by chromagenics, aquired in 24)
20	Rudolf Magnus Instituut Utrecht
21	Cyto-Barr B.V., Leiden
22	SoluCell, Espoo Finland & Gent
23	Danish University of Pharmaceutical Sciences
24	Stem Cell Innovations, Houston Texas
25	De Ruiter Seeds R&D BV, Bergschenhoek
26	Spinnovation Analytical BV, Nijmegen
27	DSM Innovation center, Eindhoven
28	Syngenta Mogen B.V.
29	Dyadic Nederland BV
30	Synthon bv, Nijmegen
31	Effecta Pharma Ltd
32	Teheran University of Medical Sciences, Iran
33	Fokker Space BV

34	TNO - Kwaliteit van Leven (Gorter Building and Gaubius Building) prevention and health
35	Flexgen
36	Chemistry
37	Galapagos
38	Biology (molecular biotechnology, molecular genetics,
39	Holland Biotechnology (HBT)
40	Drug (Leiden Amsterdam Center for Drug Research)
41	ICL-IP, Terneuzen
42	UCLA
43	Introgene
44	UMC Utrecht
45	ISA Pharmaceuticals
46	Unilever Global Foods R&D
47	JP Bioconsult, Leiden
48	University of Copenhagen
49	Sziens, Leiden
50	Vivici BV
51	LEO Pharma, Copenhagen
52	VU Medical Center
53	LUMC
54	Wageningen Universiteit
55	Merus Biopharmaceuticals, Utrecht
56	Wildcard Pharmaceuticals Consulting, Copenhagen
57	MoGen
58	ZF-Screens
59	Nijssen Light Division BV

Table 10: Overview average distances

Location	Number of organizations	Average distance
LBSP	4	0
The Netherlands	6	74,16666667
Belgium	1	134
Luxembourg	2	306
Germany	20	376,25
United Kingdom	11	389,4545455
France	9	405,2222222
Switzerland	8	571,875
Denmark	3	658,6666667
Italy	1	1108
Spain	1	1236
Israel	4	3345,5
canada	3	6733
USA	58	6822,086207
India	2	6952
South Korea	1	8633
China	1	9202
Mexico	1	9222
Japan	7	9235,714286
Taiwan	1	9502
Australia	1	16655