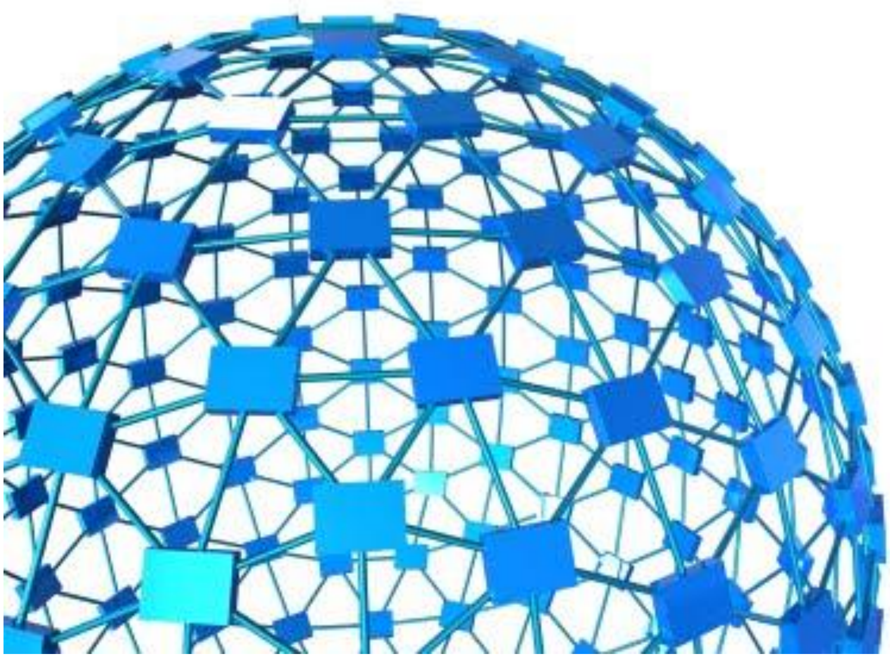
Knowledge externalities between (un)related firms

A study of technologies and labour mobility at the Leiden Bioscience Park

Gijs Janssen

Supervisor: Dr. S. Phlippen

Student number 335801
December 2010
Erasmus School of Economics, Rotterdam



ABSTRACT

This study tests whether in a region with a common knowledge base Marshallian specialization or Jacobian diversification is the best accelerator of knowledge externalities. Therefore a new measure of technological relatedness among firms and organizations is created by extending the technique of co-occurrences of technology codes in patents. This measure is calculated for firms and organizations on the Leiden Bioscience Park, and regressed against labour mobility as indicator of knowledge externality – externality because of its imperfect pricing. Labour mobility is measured by inventor mobility and mobility of members of the Board of Directors. The results indicate that inventor mobility increases when technological relatedness decreases, while Board mobility is not significantly affected by technological relatedness. This implies that when the threshold of the common science base is passed, diversification is more important for firms than specialization, while mobility at the management level is unrelated to the technologies of the firm. Inventors prefer to switch jobs between technologically less related firms, giving them more room to exploit their knowledge. The results add nuances to the general diversification versus specialization debate and give interesting insights into the role of the inventor on the park and the central role of the University of Leiden.

TABLE OF CONTENT

1. INTRODUCTION

When it comes to discussing the benefits firms can achieve by exchanging knowledge, two lines of reasoning come up: either firms will benefit most from others when their knowledge is substitutive, or the benefits are highest when the knowledge is complementary. The first argument is based on the assumption that similar technologies between firms increase the efficiency of communication and enable a high degree of specialization, while the second states that diversification leads to cross-fertilization and an alternative angle into problems that could lead to more creative results. The non-priced (externality) or unintentional (spillover) knowledge flows between firms play a major role in this discussion.

This debate has its roots in the theories of Marshall (1920) and Jacobs (1969), respectively. Empirical studies have not been unanimous so far, as the results depend strongly on the applied method and the studied industry or region. However, nuanced and detailed studies are able to formulate conditions that facilitate either of the externalities.

An important source of knowledge externalities is mobility of labour. Employees are able to share 'tacit' knowledge across firms, and form the heart of an organization. The externality increases for scientists and inventors, specialized in a certain kind of knowledge or technology. Their mobility implies the sharing of specialized, tacit knowledge that is otherwise inaccessible for a firm.

This thesis applies these theories to the Leiden Bioscience Park. This cluster of companies is built more than 25 years ago around the University of Leiden and the Leiden University Medical Centre. By now it houses more than 80 firms and organizations, all in the industry of bioscience and biotechnology.

From patent data a measure of technological relatedness is defined, both on a cluster level and on a firm level. This technological distance between firms will then be combined with labour mobility as the source of knowledge externalities. It is expected from theory that labour mobility will lead to the highest level of externalities when it takes place between the more related companies. However, as the firms on the park share a similar knowledge base, the positive externalities from relatedness are not indefinite, and might switch to diversity to avoid a technological lock-in. Although many studies have been devoted to this subject, the methodology that is applied in this study has not been used on this level and in this context yet. The ability to identify the technological distance between two or more firms gives valuable insight into the relevance of externalities for

firms, and how labour mobility is influenced by technological differences among the firms and organizations.
The research question central to this study is the following:

How is labour mobility between two firms affected by their technological distance?

The next chapter describes in detail the theoretical foundation of the research question by examining in turn the relatedness discussion of Marshall and Jacobs, discussing the theories and implications of knowledge spillovers, and finally the role of labour mobility in this theory. Chapter 3 is dedicated to a description of the Leiden Bioscience Park, followed by chapter 4 which explains the different methods of study that are used. Chapter 5 describes the empirical results in detail, the final chapter 6 will discuss these results and formulate the conclusion of this study with suggestions for further research.

2. Theoretical Framework

In this study the relationship between the technologies of a firm and its shared labour pool is examined. This chapter will elaborate on the theoretical foundations of this relationship, and creates the building blocks for the research question and the empirical study. First, agglomeration externalities and relatedness will be discussed following the arguments of Jacobs and Marshall. Second, the theory, terminology and dynamics of knowledge spillovers will be discussed, following the concept of tacit knowledge and the production function. And in the third paragraph I will discuss the importance of labour mobility for knowledge spillovers, followed by a summary and the concluding research question that will be tested in the empirical study in the following chapters.

2.1 Agglomeration externalities and relatedness

The potential benefits that can accrue to a firm caused by its location have been a major subject in economics in the recent history. Different empirical studies with seemingly conflicting results have led to a number of economic theories concerning agglomeration externalities. In this study the focus is on the agglomeration externalities that arise from the relatedness of the activities of the firms. In this paragraph I will outline the main discussion in this field of research: the Jacobian versus the Marshallian externalities.

2.1.1 Externalities from related activities

Already almost a century ago, Marshall (1920) discussed the potential benefits for firms by choosing their specific location. He found that firms would benefit more from an environment where similar firms were located, with similar skilled employees and a relatively low level of competition. Along with the identification of knowledge as being non-rival and non-excludable (hence, a potential externality) by Arrow (1962) and the endogenous growth model by Romer (1986), Glaeser, Kallal, Scheinkman and Schleifer (1992) formulated the Marshall-Arrow-Romer model, or MAR. In this study, Marshall and MAR will be used interchangeably. According to this model a concentrated industry facilitates knowledge spillovers and innovation for these firms in a certain region. Each firm in this industry specialized region can enjoy the 'local buzz'; referring to the "information and communication ecology created by face-to-face contacts, co-presence and co-location of people and firms" (Bathelt, Malmberg and Maskell, 2004, p. 33). And as externalities are defined as an effect from a certain activity to another activity which is not reflected in the costs of the former, these benefits are costless (Beaudry and Schiffauerova, 2009). Not only will these knowledge spillovers be beneficial to the firms, but it will also increase the potential of the entire region by creating a virtuous circle of regional growth and technological progress by being increasingly attractive to other firms and skilled employees (Guiliani, 2003). A more elaborate discussion on the sources, implications and empirical studies of knowledge spillovers will follow in paragraph 2.2.

Besides the localized knowledge spillovers, two other benefits accrue to firms for being co-located (Neffke, 2009):

- Labour pool. A crucial elements for firms to grow is a steady and large supply of skilled labour. By co-locating firms with similar technologies, this attracts both specialized labour and more firms to benefit from the labour pool. The role of labour mobility will be discussed in more depth in paragraph 2.3.
- Asset sharing between clients and suppliers due to close geographical proximity benefits both as it decreases both transportation as transaction costs by means of for instance face-to-face communication, exchange of parts and signing contracts.

2.1.2 EXTERNALITIES FROM DIVERSIFIED ACTIVITIES

Localization externalities originate according to Jacobs (1969) not from a specialized region, but from a diversified city. In her seminal work 'The Economies of Cities' she argues that knowledge spillovers are external to the firms' industry, and can be best served in a city with a diversified industry portfolio. The externalities from diversity are often combined with externalities originating from the urban region (Henderson, 2003), although the latter refers to benefits of a large local market and the need of a product to be consumed shortly after production (which is the case in service industries) (Neffke, 2009). In this study the focus will be on the diversification argument, although it is in practice hard to distinguish between the two.

The differences in argumentation boil down to a single, important dilemma: will the Marshallian specialization or the Jacobian diversification be the best fertilizer for agglomeration externalities?

2.1.3 MARSHALL MEETS JACOBS

Marshall and Jacobs externalities do not have to exclude each other per se. As Marshall describes the benefits of specialization in a limited geographical area, Jacobs (1969) argues that benefits from externalities can also be achieved through diversification, where cross-fertilization between industries in a city creates new ideas and innovations, which would not have come up in a region of industry specialization. This reasoning might thus explain the power of attraction of cities, and the willingness of firms to pay the relatively high rents.

Glaeser et al. (1992) test the Jacobs externalities, along with the Marshall-Arrow-Romer (MAR) externality and the argument of Porter (1990). The MAR externality implies city growth through the knowledge spillovers between firms in concentrated industries, where a local monopoly increases growth more than local competition. This externality has played a vital role in the growth of Silicon Valley, where through knowledge accumulation in a few large firms smaller firms could benefit from the knowledge spillovers. As the lack of property rights would decrease innovation incentives in the case of large externalities, monopoly power would enable the innovating firms to internalize the externalities, and thus increase more growth in the region.
Porter's argument is similar, but he argues that local competition will increase innovation and growth of the city more than a monopoly. Competition accelerates the incentives to innovate, as it is the only way to stay in business. And even more, adoption and improvement of innovations will take place at a higher rate than in a more safe and steady environment of a monopolist.
These theories are called *dynamic externalities*, as in contrast to localization and urbanization externalities, it focuses not only on the formation and specialization of cities, but also on its growth (Glaeser et al., 1992). But, as Neffke (2009, p. 28) points out, the three sources of agglomeration externalities of Marshall (labour pool, asset sharing and knowledge spillovers) already create reinforcing growth. These externalities have thus always been *dynamic*, and the distinction will therefore not be applied in the further course of this study.
The results of Glaeser et al. (1992) were not in favour of the MAR and Porter externality. A higher level of concentration of an industry did not lead to more growth. In fact, the results show a significant negative effect, which implies a confirmation of the theory of Jacobs. For the level of competition the prediction of MAR is again rejected, as the results significantly indicate a positive effect of competition on growth.
Although Glaeser et al. (1992) suggest Jacobian externalities as the main driver of city growth, it only uses employment growth as the dependent variable. In this study the level of innovativeness is the more important subject of study, as it was in other articles comparing Jacobian and Marshallian externalities (Paci and Usai, 1999b; Van Der Panne, 2004; Van Der Panne and Van Beers, 2006).

Although the empirical results of Glaeser et al. (1992) are relevant to this study, other empirical studies indicate other results. De Groot, Poot and Smit (2009) and Beaudry and Schiffauerova (2009) review many of these results and conclude that the effect of location on externalities are ambiguous, as for both specialization and diversification almost as many positive as negative effects were found. The strong difference in the results has its origins in the research methodologies that were applied, the timing of the study and its context. This indicates that definite results have not yet been found on the dynamics of knowledge spillovers and externalities among firms in general. Nevertheless, several studies show important insights in the mechanisms at work, and will therefore be discussed below.

2.1.4 Externalities and innovation

Feldman and Audretsch (1999) discuss the distinction between specialization and diversification, but do not focus on the growth of cities or regions. Rather, their focus is on innovative output, and how the composition of economic activity in an agglomeration has an effect on the potential externalities. The main conclusions are in line with the findings of Glaeser et al. (1992), as significant evidence indicates that specialization is not a promoter of innovation, but diversity is. When a common science base is taken into consideration the result still holds. This is relevant to this study as the Leiden Bioscience Park is located around the Leiden University and its Medical Centre, and the firms located on the park that are active in research and development, can all be placed under the common science base 'Biomedical', as is formulated by Feldman and Audretsch (1999). Focusing the activities within the science base thus creates less innovative output, while diversity is on the other hand beneficial for innovation. These results also hold when a firm level perspective is examined, although the common science base is of crucial importance. The study also confirms the previous findings of Glaeser et al. (1992) concerning the debate of local competition versus local monopoly; the former appears to be more conducive to innovative activity than the latter.

Van Der Panne's (2004) study is similar to Feldman and Audretsch (1999), as it also examines the effects on innovative output (in this study only on a regional level, but the results do not change when the firm level is studied (Van Der Panne and Van Beers, 2006)), but using the Netherlands as subject of study. The results contradict the previous findings of Feldman and Audretsch, as it shows that for the regional level (and the firm level) the innovative output increased in more specialized regions, where diversification shows to be negatively correlated with output. Also, the study concludes (in contrast to the results of Glaeser et al. (1992)) that competition negatively affects innovative output.

The differences in the results seem striking, but can have two important causes. First, Van Der Panne (2004) does not account for a common science base when he examines the level of diversification between firms. This already creates an important bias for his results. And secondly, the differences in the population of firms can explain the contradicting results. Not only does the number of observations differ largely between the studies (Van Der Panne studies 398 firms in the Netherlands, Feldman and Audretsch 5946 firms in the United States), but general differences in business and industry culture between The Netherlands and the United States might also resemble partly the differing results, as it did for the empirical results of Paci and Usai (1999b).

But when the study corrects for factors of time or technology, some findings can arguably be generalized. Van Der Panne and Van Beers (2006) argue that although innovators initially perform better in specialized regions when compared to diversified regions, it is also found that after the first two years of product launch the products perform better in diversified regions. As an increase in sales can imply an increase in employment, this result can be linked to the result of Glaeser et al. (1992) which was in favour of diversification. An innovating company would therefore benefit most from the specialized region in the first few years after market launch, after which it should move to a more diversified region to maximally enjoy the agglomeration externalities.
When the results are specified to technology levels, the effects of agglomeration externalities seem to become larger as the technology is getting more complicated. For instance, Paci and Usai (1999b) argue that both diversification and specialization lead to agglomeration externalities, but to a larger degree in high technology industries. Their result is based on using two measures of externalities, one for specialization and the other for diversification instead of using one measure for both. The results of Shefer and Frenkel (1998) are complementary to Paci and Usai (1999b) as it also finds a positive effect on the innovativeness for both externalities, but in this study only for high technology sectors. And Henderson (2003) (as Van Der Panne, 2004) even concludes that diversification is not present, while specialization externalities are larger for more R&D-intensive firms. Therefore, lower technology firms do not seem to reap benefits of agglomeration externalities. These results thus make a strong case for assuming agglomeration externalities in the Leiden Bioscience Park either because of specialization or diversification externalities as the firms clearly share a common science base (biotechnology).
It is crucial to note here that the causal relationship hypothesized above between specialization or diversification and agglomeration externalities can in theory also be reversed. When labour mobility is used as a measure for agglomeration externalities (as is the case in this study), labour mobility might also be the cause of technological relatedness between firms. Knowledge transfers through mobility from one firm to the other, which in the end could cause the firms to converge in terms of technological relatedness. However, for two reasons this is not expected. First, according to the resource based view of the firm the employees carry the knowledge, not the firm itself, especially when this knowledge has a ‘tacit’ nature and can therefore not be codified (see paragraph 2.3). Second, this reversed causality is only possible in a world of only a few firms influencing each other. In practice, each firm is connected to many other firms through labour mobility. An inventor

moving from one firm to another might have some effect on the firms' knowledge base, but this will also hold for all the other new employees of the firm, either coming from a firm on the cluster itself or from a firm outside of the cluster. Jousma, Scholten and Van Rossum (2009) conclude that on the Leiden Bioscience Park in 2005 75% of all employees working in companies work at a company that has an exogenous origin, which means either a division start of an already existing company or a relocation of a company external to the park. Arguing that an inventor moving in the park will have a significant effect on the technology base thus seems to be based solely on theory and is highly unlikely to persist in the Leiden Bioscience Park.

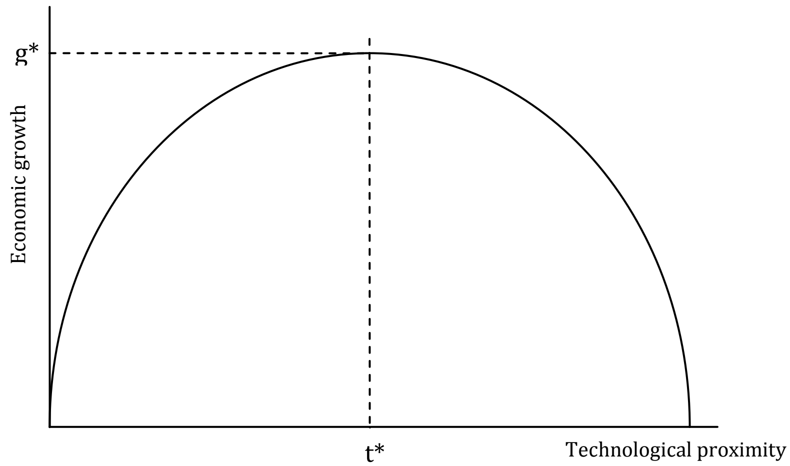
2.1.5 TECHNOLOGICAL PROXIMITY

Boschma (2005) summarizes a more theoretical discussion following the French School of Proximity Dynamics from the 1990s, where the effects of different proximities on the level of innovativeness is measured. Not only geographical proximity, or co-location is used but also the effects of different levels of organizational, social, institutional and cognitive (or technological) proximity are discussed. As in this study a measure of technological relatedness is developed (see chapter 4) I will only focus on the latter technological proximity. Geographic proximity is in this case held constant. Close technological proximity within a bounded geographical area thus refers to Marshallian specialization, as discussed in the first section of this chapter.

Next to the three benefits described by Marshall, Boschma (2005) adds benefits because of the specific characteristics of knowledge, as its 'tacit' and idiosyncratic nature requires a shared knowledge base in order to communicate, understand, absorb and process information that spills over in a spatially bounded region. However, Boschma (2005) also argues how technological proximity might have negative effects on learning and innovation whenever it is too close. It obviously lacks the benefits of diversified regions, where cross-fertilization triggers ideas and processes. This could result in a technological lock-in, where the routines of the firm limit the search for new ideas and possibilities. Access to sources of complementary knowledge and production in the rest of the world might decrease this problem, as is concluded by McCann and Simonen (2005). Boschma summarizes this trade-off in a graph with inverted u-shape, with technological relatedness on the horizontal axis and innovative performance measured vertically (see figure 1). This implies a certain optimal level of technological proximity, where a "common knowledge base is made up of diverse, but complementary knowledge sources" (Boschma, 2005, p. 64).

The Leiden Bioscience Park seems to acknowledge the lock-in problem and is constantly attracting international scientists in their patent applications and research projects (referring to their international events, see website). This has for instance led to a joint venture and settlement of a Chinese-Dutch biomedical firm on the park in May, 2010.

Figure 1: Optimal level technological proximity



Source: Boschma (2005)

The connection between the arguments of growth through diversity by Jacobs and the common science base of Feldman and Audretsch (1999) and Boschma (2005) is discussed by Frenken (2007). In his study the terminology is less about diversity and more about variety. He argues that Jacobs externalities can be measured by related variety, where a region has the highest economic growth because of a certain level of complementarity of sectors within the region. A common science base is in these concepts not enough to maximize economic growth, as complementarity is needed to avoid a lock-in. This suggests a modification of the inverted U-shape of Boschma (2005) in case of a common science base; technological variety instead of relatedness would be on the horizontal axis. At a certain level of variety economic growth is highest and optimized. A low level of variety results in a lock-in, and a high level diminishes the advantages of specialization.

2.2 KNOWLEDGE SPILLOVERS AND THE PRODUCTION FUNCTION

Knowledge spillovers has over the last few decades received a lot of attention, both from academics and spatial planners. In this paragraph the terminology and dynamics of knowledge spillovers will be discussed, followed by some important critiques by Breschi and Lissoni (2001b).
Döring and Schnellenbach (2006, p. 377) define knowledge as "all cognitions and abilities that individuals use to solve problems, make decisions and understand incoming information". This

knowledge can change with time and context and it can be spatially dispersed. Knowledge is inherently different to information, as the latter is easily codified and has a one-dimensional meaning and interpretation (for instance, the euro/dollar exchange rate). Knowledge is generally harder to codify, vague and sometimes hard to rate at its true value (Audretsch, 2003).
Whenever a recipient is new to a certain type of knowledge and is limited by its cognitive history, the probability of proper identification of the new knowledge decreases. Path dependency will therefore remain a crucial factor in knowledge interpretation (Boschma, 2005).
As knowledge becomes more specialized and complicated, it will reach a moment where it becomes uncodifiable. At that point, the knowledge 'transforms' from explicit to 'tacit', which is by definition uncodified and ill-documented, and can only be transferred by face-to-face interaction (Beaudry and Schiffauerova, 2009). Although the increasing possibilities of interaction through the internet and other social media seem to decrease the necessity of geographical proximity for the interacting actors (Vacaro, Veloso and Brusoni, 2009), frequent interactions where knowledge can be learned through action and reaction will remain more efficient. More formally: "the marginal costs of transmitting knowledge, especially tacit knowledge, is lowest with frequent social interaction, observation and communication" (Audretsch and Feldman, 2003, p. 7). Von Hippel (1994) classifies knowledge into 'sticky' or 'non-sticky'. Knowledge is defined 'sticky' when it is costly to acquire, transfer and use. Combining this with the definition of tacit knowledge, the larger part of the 'sticky' knowledge will be tacit. Especially in R&D-intensive and high-tech industries, tacit and 'sticky' knowledge can be very important.

Knowledge is an input for firms, but different from the other inputs labour and capital, as the stock of knowledge does not decrease when it is used. However, it does increase in a cumulative way when more knowledge is created, for instance within a firm by research and development (Boschma, 2004), or by universities and knowledge institutes.
Firms invest in new knowledge and technologies, but generally can only appropriate a portion of the investment. Part of the investment will flow out of the firm by means of externalities. As employees of these firms have by definition limited cognitive capabilities, they can only see a fraction of the possibilities of their newly created technologies. By discussing the innovations with others they can be inspired to new applications of the innovation, and internalize a larger part of the initial investment. However, by doing this employees unintentionally spill over knowledge to others, which is acquired for free, and might be used in the others own business routine. When this process is repeated on a larger, but geographically bounded scale, firms benefit to such a large

degree from the knowledge spillovers of others, and get an even larger return on their investment than they would have gotten when they had fully internalized their expenditure.

This line of reasoning is summarized and criticized by Breschi and Lissoni (2001b). They formulate it as 'the three-step logical chain':

1. Knowledge from firms and/or universities is transmitted to other firms
2. The knowledge that spills over is a pure public good (Arrow, 1962), characterized by its non-excludability and non-rivalry
3. However, as knowledge is largely 'tacit', the knowledge is context dependent and hard to codify, hence more easily transmitted through face-to-face interaction. This requires firms to co-locate, which means that the 'tacit' knowledge is only a public good on that location.

The main argument of Breschi and Lissoni (2001b) is in the use of the term 'localized knowledge spillovers' and the proxies that have been used in literature to measure them. Although the potential importance of these spillovers is not denied, they do argue that most studies use measures of pecuniary externalities to draw conclusions about 'pure' knowledge externalities, while in practice knowledge spillovers are generally not 'involuntarily', but regulated by firms to enhance the appropriability of their innovations. According to Breschi and Lissoni (2001b) the 'localized knowledge spillovers' as a concept has been largely abused in academic literature, and any attempt to measure it has to be carefully defined and categorized. In this study labour mobility will not be defined as a localized knowledge spillover, but as a localized knowledge externality, similar to the agglomeration externality as it is defined by Marshall (1920) but focused on knowledge. Where firms may enjoy the expertise of specialized knowledge workers moving around in the Leiden Bioscience Park, it is not assumed that this process is unintentionally or without costs. However, it is assumed that firms do not pay the entire compensation for the received externality, and the sum that is paid will predominantly go to the knowledge worker, and not to the firm where the employee comes from (Møen, 2005). If hypothetically this was the case, the receiving firm would have to pay to all past employers of the moving knowledge worker as a compensation for the knowledge received. Ironically, this does happen in soccer, where in case of a transfer of a player, the buying club is obligated to pay at least 5% of the total sum of transfer to clubs where the player use to play until he is 23 years old (FIFA, 2003). The role of labour mobility and tacit knowledge will be discussed in more depth in paragraph 2.3.

2.2.1 Source of localized knowledge externalities

As was discussed in paragraph 2.1, a long debate has been going on about which business environment promotes agglomeration externalities the most. But whether diversification or specialization is the best condition cannot be stated without knowing the circumstances. The empirical results fluctuate substantially, and the differences in outcomes are mostly related to the different methodologies applied.

Nevertheless, considering the Leiden Bioscience Park as a high-tech and R&D intensive cluster, it can be assumed that the companies benefit more from being located in a specialized business park than in a more diversified urban location. In theory these benefits are described by Marshall (1920) and later formalized by Glaeser et al. (1992) as asset sharing, a common labour pool and knowledge spillovers. These components however remain theoretical, and have shown to be hard to quantify. In this section the focus will be on the production function, how this can explain knowledge externalities and what the corresponding empirical results conclude.

In economic theory, the production function of the firm by Solow (1957) is an important cornerstone in disentangling firm behaviour. For the production of innovations and technological change the function was expanded to include the production of knowledge as an input for innovative output (Griliches, 1979), where later R&D was defined as the primary indicator of knowledge inputs (Cohen and Klepper, 1992). The knowledge production function is formulated as follows (Audretsch and Feldman, 2003):

$$I_i = \alpha RD_i^{\beta} HK_i^{\gamma} \varepsilon_i$$

Where I is the degree of innovative activity, explained by the inputs RD as R&D and HK as human capital. The i stands for the unit of observation, for instance countries, industries or firms. Although the reasoning is straightforward (innovative output is a function of innovative inputs), empirical studies have not been consistent with this function. When it is tested for countries or industries the results are robust, but studies for the relationship among firms with different sizes show other results. As this knowledge production function only holds at the aggregate level and not on the firm level, this may imply the presence of externalities (Audretsch and Feldman, 2003).

The models of these externalities were already described and discussed by Marshall (1920) and Jacobs (1969) as was described in paragraph 2.1. Although the specialization versus diversification discussion is still subject of studies, the mere presence of externalities and knowledge spillovers within a certain spatial agglomeration was theoretically evident. On that assumption, Jaffe (1989) formulated a new knowledge production function, including spatial and product dimension (re-formulated by Audretsch and Feldman, 2003):

$$I_{si} = \alpha IRD^{\beta_1} * UR_{si}^{\beta_2} * \left(UR_{si} * GC_{si}^{\beta_3}\right) * \varepsilon_{si}$$

Again, *I* is innovative output, *IRD* is industry R&D, *UR* is academic research expenditure and *GC* is a measure of the geographic coincidence of university and industrial research activity within a state. The small *s* and *i* represent the unit of observation, respectively the state and the industry. With this extension, Jaffe (1989) lifted the unit of observation from the firm level to the spatial level, allowing for localized knowledge spillovers and externalities to occur.

Jaffe, Trajtenberg and Henderson (1993) test the above knowledge production function by taking geographic location of patent citations as a proxy for knowledge spillovers. By using these citations they make an attempt at opening the black box of knowledge spillovers. Although Krugman (1991) argues that knowledge flows are invisible by definition, Jaffe et al. (1993) do see a paper trail of knowledge by means of patent citations. And as these citations can be localized, so can the patent citations and thereby the knowledge spillovers.

Their results are in line with the theory. The paper trails of the patents citations are geographically localized, which means that patents are more likely to be cited by others located in relative close geographical proximity than by actors located further away. However, when is corrected for technology classes, the positive effect disappears. When a patent is cited by another from the same classification, the probability of it origination from the same geographical location is not higher than when it is cited from another classification. As Jaffe et al. (1993) only use the 'primary' patent class this result may be biased and can change when a more extensive measure of technology classes is used. Although in this article I do not test for knowledge spillovers (they are assumed), I do use a more elaborate method of defining technology classes and measuring technology differences between firms.

The effect of technological proximity is also subject of study of Autant-Bernard (2001). She concludes even stronger than Jaffe et al. (1993) that firms only enjoy knowledge spillovers whenever they are co-located, but when technological proximity is taken into account these effects are significantly smaller. Spillovers depend almost entirely on the labour pool of researchers in the area, despite of the conceptual critiques of Breschi and Lissoni (2001b). This study does emphasize the role of labour mobility, but uses technological and geographical distance as complementary concepts. This study takes geographical proximity as a constant and will therefore create a more reliable conclusion of the effect of technological proximity on labour mobility as source of externalities. The role of labour mobility will be discussed extensively in the next paragraph.

2.3 Role of labour mobility

When an idea is born, it has no material content. Therefore, it is virtually unlimited in space, not bounded by spatial restrictions or availability of resources. However, to enable practical use of the idea, it has to be transferred by means of communication. And although modern communication technologies make the transfer of ideas more easily across regions and countries, "intellectual breakthroughs must cross hallways and streets more easily than oceans and continents" (Glaeser et al., 1992, p. 1127). Therefore, ideas are best appropriated in spatially bounded regions, where communication is abundant and its quality is high. And when the idea or knowledge is highly complex, context-specific and uncodifiable, it becomes 'tacit' and only transferrable through face-to-face interaction (Almeida and Kogut, 1999). The importance and relevance of inter-firm interaction and labour mobility is evident, especially considering the vital role of labour in the firm according to the classic resource based view. In this paragraph, the role of labour mobility will be further emphasized using theoretical arguments and empirical evidence, justifying labour mobility data as the primary source of knowledge flows in the Leiden Bioscience Park.

Tacit knowledge is subject of discussion since it was first introduced by Polanyi (1966) as a philosophical concept. He argues that people are able to know certain things, without being able to formulate this knowledge (for instance in face recognition: we can identify a familiar face out of thousands, but it is hard to describe or draw that particular face). This tacit knowing can also be taught in a way that one can learn something without knowing what one has learned. "One can know more than one can tell" (Polanyi, 1966, p. 4) Learning this tacit knowledge goes subconsciously, but best when the contact is direct and face-to-face. When this concept is translated to the social-economic science, we can thus argue that through face-to-face learning more knowledge can be learned than was practically transferred. Therefore, the concept of tacit knowledge plays a major role in knowledge spillover theory. Tacit knowledge is abstract, created by personal experience and difficult to transfer over larger distances. It is only understood by people who have had the same personal experience with the knowledge and share a common social context (Wilson and Spoehr, 2010; Breschi and Lissoni, 2001b). These conditions restrict the transfer of tacit knowledge inevitably to a geographically bounded location. We have already seen the successful attempt of Jaffe et al. (1993) in measuring this tacit knowledge, in this paragraph the focus will be on labour mobility as the means to disperse and diffuse tacit knowledge.

Breschi and Lissoni (2001b) discuss two ways tacit knowledge can diffuse in a geographically bounded region. The first one lies in the nature of tacit knowledge itself, as it generally involves a language or codebook of its own. Only the member of this epistemic community that are familiar with this knowledge, and therefore with its language, can decide upon sharing this knowledge with outsiders. This language thus acts as an 'exclusionary device' for others, even people located in the same spatially bounded area. Although Breschi and Lissoni (2001b) argue that this mechanism enables tacit knowledge to be exchanged over larger distances provided that both actors are familiar with the specific language, this is not assumed here as Polanyi (1966) specifically mentions the physical proximity as a condition for the exchange of tacit knowledge.

The second mechanism of tacit knowledge diffusion is by labour mobility. In academic literature labour mobility has been interchangeably called a form of knowledge spillover (Almeida and Kogut, 1999; Balsvik, 2006), an agglomeration externality (Marshall, 1920) or knowledge transfer (Zucker, Darby and Armstrong, 1998; Breschi and Lissoni, 2001b). The distinction is not always clear with respect to pecuniary compensation, and depends to a large extent on the type of labour that moves between firms and/or universities. In this study I apply labour mobility as a knowledge externality where knowledge is transferred, but with inefficiencies. I assume that mobile scientists receive some form of compensation for the knowledge they carry and bring to the firm, but this compensation will not cover exactly the benefits that accrue to the labour-receiving firm. Appropriability of knowledge, especially when its content is still unknown, can never be appreciated perfectly. Furthermore, the inventor pays for the knowledge that is learnt within a firm through lower wages earlier in their career, and higher wages in a later stage. This resembles some form of internalization of the potential externalities associated with labour mobility (Møen, 2005). Hence, it is worth emphasizing that in this study labour mobility is interpreted as a transfer of tacit knowledge, and therefore by definition imperfect in setting its price. This does not rule out the presence of a 'local buzz', defined by Bathelt et al. (2004), as knowledge may still spill over due to communication and interaction within the geographic region. The measurement of a 'local buzz' would require other data and other methods than the ones applied here, and will not be tested empirically. A more qualitative study could shed more light on the presence of such a local buzz in the Leiden Bioscience Park.

The diffusion of knowledge through labour mobility is generally measured by focusing on the mobility of the star-scientists (Zucker et al., 1998), technical personnel (Møen, 2005) or inventors (Ibrahim, Fallah and Reilly, 2009; Agrawal, Cockburn and McHale, 2006), but as these groups by and large coincide their implications can be generalized. And as knowledge diffuses more quickly between co-located actors because of the lower communication costs, higher likelihood of interaction through chance and higher likelihood of social relationships (Agrawal et al., 2006), labour mobility plays an important role in localized knowledge diffusion. Moreover, the network of actors is found to be the primary source of knowledge (Kogut, 2000); in the labour market of inventors learning has proven to be the primary reason behind hiring (Palomeras and Melero, 2010); and the geographic extent of a knowledge spillover is almost completely controlled by the inventors (Breschi and Lissoni, 2006). Furthermore, as formal relationships between firms do not appear to be strongly localized, the local labour market might be the crucial link to localized growth (Arita and McCann, 2000). Not only is human capital the most important channel of knowledge diffusion (Breschi and Lissoni, 2001a), it is argued to be of greater importance to the firms than R&D expenditure (Autant-Bernard, 2001).

These findings and theories create a strong argument to use labour mobility as an important source of knowledge diffusion in the Leiden Bioscience Park. As is explained in more depth in chapter 4, two (complementary) forms of labour mobility are used. The former one is mobility of inventors within the park, based on patent information. The latter one is the mobility of member of the Board of Directors of the firms in the park. These data are provided by the research institute at the park, and comprise all current members of the Boards, and their work history at the park. These forms of mobility are used both separately and as a common group. It can be expected that inventors carry more tacit knowledge than members of the Board, but measuring the differences in this perspective lies outside the scope of this research. Nevertheless, differences in the empirical results among the groups are discussed in chapter 5.

2.4 RESEARCH QUESTION

To conclude the above theoretical discussion, I shall shortly summarize the different arguments, and explain how the theory leads to the research question which is tested in the remainder of this thesis.

The discussion concerning the best conditions for firms to benefit from agglomeration externalities has two major camps, and has not been resolved despite many empirical studies. The diversification argument of Jacobs (1969) and Marshall's (1920) argumentation of specialization have empirically shown to be rather balanced (De Groot et al., 2009), as results seem to depend heavily on many other factors. I have shown that for specialized, high-tech clusters (such as the LBSP) specialization is the best facilitator of agglomeration externalities, which I therefore assume to be present in the case study. This level of specialization can create a potential lock-in, but also creates benefits as firms can more easily understand, absorb and process information from others because of their shared knowledge base (Boschma, 2005). The “technological distance” between firms is formulated in the chapters 4 and 5, and is based upon knowledge produced by the firms in the form of patents.

From the agglomeration benefits according to Marshall, which are localized knowledge spillovers, asset sharing and a common labour pool, I go into the former in more depth, as this remains a black box in academic studies (Breschi and Lissoni, 2001a). As involuntary spillovers mainly comprise tacit knowledge , this has to be transferred, which due to the nature of tacit knowledge can only be done through face-to-face interaction between people. Although communication and personal relations will play a role in the dispersion of tacit knowledge, the main driver will be labour mobility.

In the empirical study I will test the combination of the technological distance of the firms on the park, and the mobility of labour between the firms. The question I will answer is the following:

How is labour mobility between two firms affected by their technological distance?

The theoretical discussion suggests that labour mobility will be larger between firms with related technologies, but up to a certain degree where spillovers are at an optimal level (as is suggested by Boschma, 2005). An examination of the Leiden Bioscience Park (see chapter 3) will give additional insights into the relatedness of the firms at the park, and whether this relatedness is at a certain level that firms either look for more or less related technologies (Frenken, 2007). As the method applied creates an objective and strong relative measure of technology (see section 4.1.3) it is expected that the results are even stronger than theory suggests. This will be tested extensively in the following chapters. For labour mobility two complementary measures will be tested and compared (inventor mobility and mobility of members of the present Boards of Directors of the firms). Furthermore, technological relatedness is examined in more depth. A technology map of the park will be created, indicating which technologies are most prominent and what the dispersion of the technologies looks like. These and other measures will be calculated and compared over time to see how technology has changed in the more than 25 years of history.

The final conclusion of this thesis compares the results of the empirical study with the theoretical expectations formulated in this chapter. It will discuss the implications from the perspective of the entire region, of the individual organization and the employee (either inventor or Board member).

3. LEIDEN BIOSCIENCE PARK

In this chapter I will discuss the case that is going to be studied. I will shortly go through the history and nature of the park, I will discuss the triple-helix system and how this is adopted on the park, and finally I will go into the growth of the park over the years.

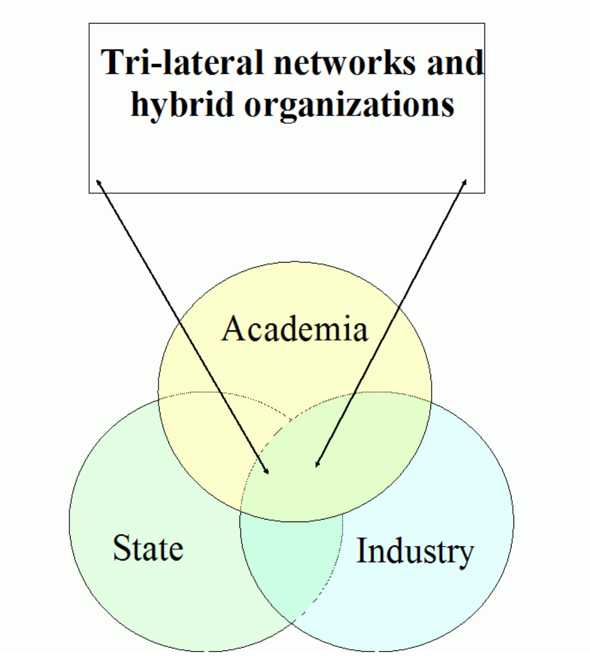
The Leiden Bioscience Park (LBSP) was established in 1984, as a collaboration of the City Council of Leiden and Leiden University. Both parties believed in the potential of bioscience, and anticipated a science park which would enhance economic growth of the entire region. The park was located next to the Leiden University Medical Centre (LUMC), and started with 3 organizations: TNO, the Dutch public research organization which would focus on life sciences, Leiden University and the LUMC.

Biotechnology or bioscience is "the use of (parts of) organisms for the development and production of new products and technologies" (definition website of the park), for instance to grow better crops, create better medication and drugs, or improve the quality of food. In biotechnology three specializations can be differentiated: 'green', related to agriculture and food; 'white', representing biotechnology for industrial processes; and 'red' biotechnology, dedicated for medical solutions and life sciences. The LBSP comprises the latter form of bioscience.

When the park was founded in 1984, it followed the so-called 'triple helix' model. In this model, balanced and dynamic interactions between government, universities and the industry result in profitable business and an effective knowledge cluster. The essential feature of the triple helix is the overlap between the three actors. In these institutional spheres, hybrid organizations develop to form a dynamic link between the different actors (Etkowitz, 2002). See figure 2 for a graphic illustration.

These institutional, overlapping spheres between the different actors can take many forms and work in many reciprocal ways. Local government and new entrepreneurs are brought together in one of the two incubator-buildings, where starting businesses are supported with the necessary facilities. Closely linked to this is the Technology Transfer Office, which enables commercialization of promising research, both industrial as academic. And the multiple collaborations between the industry and the Leiden University and the LUMC in filing patents underlines the triple helix concept on the park.

Figure 2: The triple helix model



In the last 25 years, the LBSP grew from the 3 collaborating parties to a total of 87 organizations in 2008; 75 private firms and 12 other organizations (either non-profit or research and educational). Throughout the years, 94 firms have entered the park, mostly by start-ups or spin-offs. The prominent role of the LUMC and Leiden University in the park is exemplified by the fact that all but 3 of the 34 spin-offs involved at least one of these organizations. A more extensive table can be found in the appendix (table 10) (Jousma et al., 2009).
The number of people employed at the park almost doubled between 1985 and 2005 from 5108 to 9936, more than half of that can be allocated to the companies. But still over 70% work in public education and research. The extensive table 11 can be found in the appendix (Jousma et al., 2009). These growth figures can have two major implications for the network analysis of the technologies on the park in section 4.1.2: either the number of patents increased over time and the focus on the technologies remain, or the increase of patents have led to a broader set of technologies on the park. The converging or diverging network relates closely to the cognitive proximity discussion of Boschma (2005) and Frenken (2001), and will be discussed in chapter 6. An in depth network analysis of the pharmaceutical industry by Orsenigo, Pammolli and Riccaboni (2001) identifies the dynamics of early entrants enjoying first mover advantages and possibilities to specialize their initial general knowledge, while already specialized incumbents face difficulties in absorbing the new general knowledge in the network. Although testing this for the LBSP is beyond the scope of this study, it might help in explaining some of the results, especially since the differences between

the actors in the LBSP can be substantial; where the University has a broad and more fundamental research base, a private firm is generally more downstream in its knowledge production.
As the role of the University, research centres and other non-profit organizations are not at the centre of this study they are not distinguished from private firms and are all labelled as firms, unless it stated otherwise. In chapter 6 the role of the University is given explicit attention.

4. METHODOLOGY

The theoretical analysis in chapter 2 supported the need to examine in detail the relationship between technological relatedness and knowledge spillovers, indicated by labour mobility. In this chapter I will go into the data that is used and the methods applied.
Recall the research question formulated in chapter 2:

How is labour mobility between two firms affected by their technological distance?

This research question comprises two parts: labour mobility, and the technological distance. For both parts a different dataset and source will be used and eventually combined in the regressions. In paragraph 4.1 I will go into the data used for measuring technological distance and inventor mobility, paragraph 4.2 is dedicated to another form of labour mobility which is the mobility of member of the Boards of Directors of the firms on the park. And the final methods that are used for the regressions are explained in paragraph 4.3.
Technological distance, proximity or relatedness all refer to the distance between firms in terms of the technologies they use. When firms are active in the same industry, they are technologically closer to each other compared to firms from different industries. And according to theory discussed in chapter 2, this distance can have a significant impact on the firms' innovation and growth potential. But as it might be straightforward to see how the technological distance increases when the industry of the firms is no longer the same, it gets more difficult when firms in two already differing industries are compared, and how the distance within an industry is determined. Add to that the problem of identifying the firms' technologies, and it becomes clear that measuring technological distance is not a walk in the park.
The key to measuring technological distance is patents. Patents contain a large amount of information, are updated frequently and accessible for everyone. The industry of the Leiden Bioscience Park (LBSP) (bioscience) is beneficial to the use of patents, as patents play a significant and important role in the creation and dissemination of knowledge (Owen-Smith and Powell, 2004).Two parts of its information are essential for this study: the location of the filing of the patent, in this case the LBSP, and the technological codes on the patent which will be used to measure technological distance.
To capture all the patents filed at the LBSP I use the OECD REGPAT Database. This is a regionalized patent database, based on the European Patent Office's (EPO) *Worldwide Statistical Patent Database*

(PATSTAT) and the *OECD Patent Database*. Each applicant and inventor is categorized by their addresses into certain regions. These regions are primarily based on postal codes or town names.

4.1 TECHNOLOGICAL DISTANCE

The REGPAT database actually comprises two dataset: granted patent applications filed to the EPO, and granted patent applications filed under the international Patent Co-operation Treaty, both ranging from 1977 to 2007 for priority data. In this study the former dataset will be used, as most patents are at least filed in Europe (and some also internationally).

The dataset gives an overview of two key elements: the names and locations (on a NUTS-3 level) of the applicants and inventors, and the IPC-codes (International Patent Classification) per patent, indicating the technology that is used. Furthermore, the priority and application year of the patents are provided, which make it possible to identify changes over time. Table 12 in the appendix gives an overview of the tables as they are in the raw data.

As the raw dataset comprises all patents around the world from 1977 to 2007 (priority year) the first step is to select the relevant patents. In this study I want to measure the technologies present at the park, which implies that I need to collect all patents that are created on the park or in collaboration with the park since it was founded in 1984. Unfortunately, a patent does not say where it was created. Therefore several steps will have to be taken in order to extract the right dataset from the raw data.

- First, an obvious selection can be made based on the locations of the firms on the patents. The REGPAT database' geographical area closest to the LBSP is called 'Leiden & Bollenstreek' (NUTS-3 code NL331), and comprises the entire municipality of Leiden and some surrounding villages. As the LBSP is the only higher-technology area in this region, most of these patents belong to the LBSP. A manual walk-through of the company names resulted in a list of 517 patents which are created by firms on the LBSP.
- Because in this list the Leiden University is mostly used as a single entity, it applied a relatively large number of patents over the years. Although it is from this data not possible to distinguish different faculties or departments of the University, it is possible to differentiate between the University of Leiden and the Leiden University Medical Centre (LUMC). In most cases the accurate name was already provided, and in cases of doubt the address made the distinction clear.
- But this list does not necessarily comprise all patents created on the park. As there are also some subsidiaries located on the park from firms outside the region, their patents might be applied for

by their headquarters, located elsewhere. To tackle this the firms that are or have been located on the park is compared with the list created in the first step. Then the firms that are not in this list are run through the large raw dataset, to see whether they applied patents at all. When they did, I checked whether Dutch inventors worked on their patents. At this moment only a few firms (Genencor and Centocore) and several patents (164) remain. As the inventors of the patents are Dutch, I assume (at least part of) the patent is created in a subsidiary in The Netherlands, and as all these subsidiaries are located on the LBSP I assume that the knowledge the patent holds is created and present in that particular subsidiary. These patents where obtained using the Dutch Patent database Espacenet and are therefore in some instances more recent than the patents from the REGPAT database. An overview of the applied patents per year is provided in table 13, and will be discussed in paragraph 5.1.

- Finally, the Dutch technical research institute TNO is also present at the park, but all their patents are applied for by their headquarters either in Delft or The Hague. An obvious link between a TNO-patent and the LBSP is missing, although from the information on their website it is clear that research is taking place at their subsidiary on the park. Therefore I contacted TNO in person, and kindly received a list of 39 patents created on the park since their settlement on the park in 1991.

The further processing of this dataset is in three directions. First, the labour mobility of inventors on the LBSP is extracted. Secondly, I use the International Patent Classification (IPC) codes on the patents to measure the technological distance between technologies, which will be used to create a technology map of the park. Using this in combination with the year of patent application, it is possible to observe the technology map over time, and identify changes of the most central technologies and the density and diversity of the technologies. And third, the IPC codes will be used to create a map of technologies of the firms, and identify technological distance between them. The method applied in this case will be explained in section 4.1.2.

4.1.1 Inventor mobility

Patents always hold information about the inventors. The REGPAT database provides a unique person id, along with the address of the inventor at the time of filing. Whenever an inventor files patents under a different applicant (the firm that filed the patent), I consider this a movement of labour. However, as the data does not provide information about the exact time of movement but only about the length of time between two patent filings, it is not possible to observe movement over time. Furthermore, only inventors that patented their inventions are taken into account; more inventor movement on the park can be expected, the movements of inventors in this study can be the tip of the iceberg.

By building matrices of inventors working at one or more firms, I can create a firm by firm matrix based on co-occurrences of inventors. This matrix can be translated into a map of firms at the park, which are connected whenever they share one or more inventors.

However, labour mobility might be harder to identify in two specific situations: whenever an inventor applies for two patents for two applicants in the same year, and when an inventor applies for a patent with multiple applicants. In the first situation it is not possible to identify the direction of the movement; this movement will therefore be interpreted as bi-directional, so both firms receive and send this knowledge. In the second situation it is not possible to identify to which firm the inventor belongs. However, as it is the objective of this study to identify knowledge externalities, this can be seen as a pure knowledge externality (which might go via the inventors that developed the patent), and will therefore also be interpreted as a bi-directional connection between the two (or more) applicants.

4.1.2 Technology maps

Technological relatedness describes how far technologies are different or similar from each other. Engelsman and Van Raan (1991) describe a method based on patents, more specifically the IPC-codes on the patents. IPC stands for International Patent Classification, developed under the 1971 Strasbourg Agreement. These codes are formulated and updated regularly by a Committee of Experts, consisting of representatives of countries that signed the Agreement and observers from other organizations, such as the more local patent organizations as the EPO or JPO (Japan Patent Office). These classifications consist of 5 parts or steps, each further step defining a more detailed level of classification (WIPO, 2009).

A	**61**	**K**	**61**	**/**	**01**
Section	Class	Subclass	Main group		Sub group

As these classifications are highly detailed, each patent holds one or more IPC-codes, and in some instances even more than 20 codes.

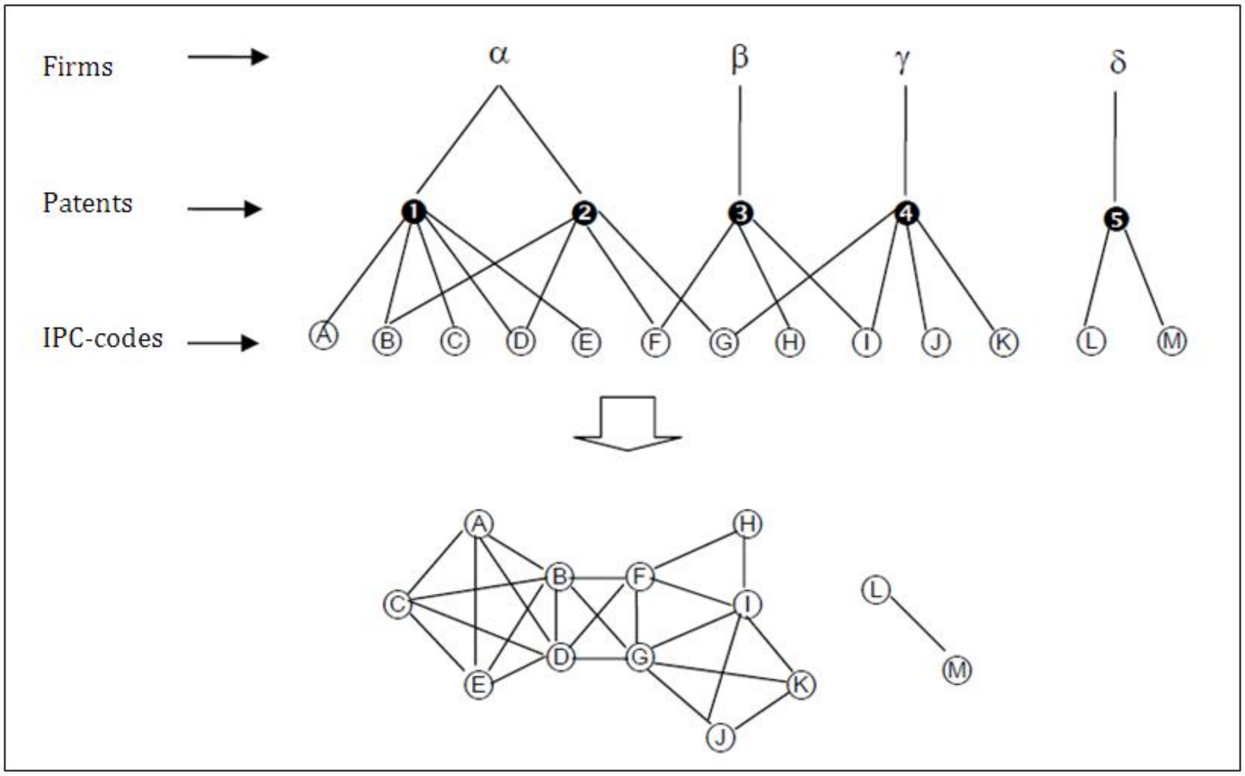
From these technology classifications it is possible to identify a relative distance between each technology. This method, extensively described by Engelsman and Van Raan (1991) uses co-occurrences of IPC-codes in different patents as a way to test how related technologies are. Whenever two IPC-codes co-occur together in several patents, it is assumed these technologies are relatively strong related; when IPC-codes co-occur only via one or more other IPC-codes, they have a relative weak relation.

The first step in this method is identifying the relevant IPC-codes. As described in the section above I selected all patents that where either developed on the LBSP or developed elsewhere in cooperation with inventors or firms located on the park. In these data a patent can have a certain number of IPC-codes ranging from 1 to a maximum of 35 IPC-codes per patent. I decided not to use the entire IPC-code, but only the first 4 parts. Although I am aware that taking not the entire code might make the relatedness in the park seem stronger than it would be when all 5 parts would be used, this study only focuses on the internal relatedness of technologies, and does not compare with other clusters. Next to that, the technology specification level is already very sophisticated with the first four parts, and on a more practical note, it increases the readability of the network pictures in the following chapter.

In the next step a 2-mode matrix is created, with the unique patents vertically and unique IPC-codes horizontally. A 1 indicates a presence of that particular code in that particular patent, and a 0 otherwise.. From this table it is already possible to identify IPC-codes that are more co-occurring than others, but to get a better overview of co-occurrences it is necessary to create a 1-mode matrix.

The step of going from a 2-mode to a 1-mode matrix can best be illustrated by using a figure from Breschi and Lissoni (2006). The figure below (figure 3) is a simplification of the data that is used in this study, and represents firms owning one or more patents with each patent having several IPC-codes. As some IPC-codes seem to be part of more than one patent (B and D are both part of patent 1 and 2, G belongs to 2 and 4, etc.), these IPC's thus *co-occur* in patents and are relatively related. The lower graph in figure 1 indicates the connections between the IPC-codes based on their co-occurrences. This is the 1-mode graph that followed from the 2-mode data above.

Figure 3: Co-occurrences IPC-codes



Source: Breschi and Lissoni (2006)

This 1-mode matrix has both on the vertical and the horizontal axis the unique IPC-codes. For each patent holding more than one IPC-code, each combination of IPC-codes is indicated in this matrix with a 1 (or added with 1 if a co-occurrence already existed). This creates a symmetrical matrix, which can be used as an input for different kinds of network graphs and calculations. In this case I focus on the development of technologies on the park over time, using time cohorts of 5 years from 1984 onwards to examine how the relatedness of technologies evolved through the years, and how both centrality of the different technologies and centralization of the entire network of technologies have changed over time.

Degree centrality is the measure of centrality applied, giving an indication of the distance between the centre of the network and the periphery (Wasserman and Faust, 1994). When a certain technology is related to many other technologies directly it has a high level of degree centrality. Degree centrality is measured on three different levels: micro, meso and macro level. The micro-level is the degree of each IPC-code in the graph, the meso-level is the average degree and standard deviation of all IPC-codes, and the macro-level is a total measure of centralization for the entire network.

To enable comparison of degree levels among graphs it is vital to use normalized degrees. As UCInet can only calculate normalized degrees for binary data the normalized degrees have to be calculated using the primary 2-mode data of patents and their IPC-codes (the 1-mode network contains valued data). Borgatti and Everett (2005) describe a method where normalized degree can be calculated without losing information about the size of the connection between the IPC-codes. Normalized degree centrality is calculated by dividing the nominal degree of an IPC-code with the number of patents in this network.

The meso-level of degree centrality is represented by the average normalized degree and its standard deviation. Changes of this ratio give information about the distribution of the degrees among the IPC-codes of the network.

Measuring centralization of the entire network (macro-level) can follow the same formula that Freeman proposes in his seminal paper of 1979, where the sum of the differences between the most central actor and the other actors is normalized by dividing it by the maximum degrees over all connections. See the next equation:

$$\frac{\sum(c_* - c_i)}{max\sum(c_* - c_i)}$$

Where c_* is the highest level of normalized degree centrality and c_i the degree of every other actor (or in this case IPC-code). Calculating the maximum as is denoted in the denominator can be done with the next equation. The n_0 represents the number of IPC-codes, n_1 the number of patents in this network.

$$\frac{(n_0 n_1 - n_1 - n_0 + 1)(n_0 + n_1)}{n_0 n_1}$$

Multiplying the results with 100 gives the centralities in percentages.

The results of these tests and the accompanying network maps can be found in section 5.1.2.

4.1.3 Firm Relatedness

The distance between technologies, calculated according to the method above, does not give any information about the technological relatedness of the firms on the park. To calculate this a few more steps need to be made.

First, the geodesic distance between the different technologies needs to be calculated. This distance is in fact the formal representation of the network maps that can be created using the 1-mode matrices. The distance that needs to be covered to go from one technology to another (either directly, which makes the distance to be 1, or via other technologies, which increases the distance above 1) is the geodesic distance. The network analysis program UCInet (Borgatti, Everett and Freeman, 2002) can do this procedure immediately.

From the network map of the technologies (see the empirical results in the next chapter) it can be seen that not all technologies are connected to each other, in some cases not even via other technologies. This is similar to the simplified example in figure 3, where the IPC-codes of patent 5 are not connected to the others. As with figure 3, figure 10 shows that there is a clear main component in the graph where some technologies are not connected to. As this implies that there is an infinitely large distance between the unconnected technologies, this cannot be used in the further study. The technologies not connected to the main component will therefore be discarded; in total this involves 37 technologies (of the total of 224). Not only will these technologies not be used, but the patents that hold these technologies (like patent 5 in figure 3 does) can also not be used. In total this involves 25 patents, owned by 7 different organizations. The full list of excluded patents and their proprietors is included as table 13 in the appendix. As all patents of both Dutch Space B.V. (11 patents) and Produvation B.V. (1 patent) are unconnected to the main component, these firms will not be used in the further course of this study. The remaining 187 IPC-codes (of 36 firms) will be used as follows.

The geodesic distance between the technologies have to be transferred to a certain (relative) technological distance between firms. In order to do that, I have to transform it to a patent-level first. I do this by taking the geodesic distances of all IPC-codes of each patent to all IPC-codes of each other patent. Taking the weighted average of these geodesic distance creates a relative technological distance between all patents. This normalizes the outcomes, so that differences in the number of IPC-codes no longer affect the results. More formally:

$$TechDistPat_{i,j} = \frac{\sum\left(GeoDist\left(IPC_{x,j} \rightarrow IPC_{y,j}\right)\right)}{\left(IPC_{N,i} + IPC_{N,j}\right)}$$

Where $TechDistPat_{i,j}$ stands for the technological distance between patents *i* and *j*, $IPC_{x,i}$ and $IPC_{y,j}$ is the IPC-code *x* of patent *i* and *y* of patent *j* respectively, and $IPC_{N,i}$ and $IPC_{N,j}$ represent the number of IPC-codes of patent *i* and *j* respectively.

The patents can directly be linked to the firms that own them, and this is done in a similar manner as is explained above. I use the technological distance between the patents, and take a weighted average of the distance between the patents of two firms to normalize the results again. This causes the number of patents a firm has not to influence the outcomes, and may also acts as a means to control for firm size (assuming larger firms own more patents than smaller firms). The formal equation has the following form:

$$TechDistFirms_{k,j} = \frac{\sum\left(GeoDist\left(Patent_{d,k} \rightarrow Patent_{e,l}\right)\right)}{\left(Patent_{N,k}Patent_{N,l}\right)}$$

Where $TechDistFirms_{k,l}$ stands for the technological distance between firms *k* and *l*, $Patent_{d,k}$ and $Patent_{e,l}$ is patent *d* of firm *k* and *e* of firm *l* respectively, and $Patent_{N,k}$ and $Patent_{N,l}$ represent the number of patents of firm *k* and *l* respectively. The resulting firm by firm matrix thus represents a certain level of technological distance between each firm. These numbers are not directly used in the regression (see paragraph 4.3) as its interpretation is not straightforward. Therefore the natural logarithm will be computed of each distance to enable a log-linear interpretation of the regression results.

Next to the application of the technological distance in this study, it can also be of direct relevance for firms and consultants in their search of technological progress. As firms look for either complementary or more specialized knowledge a direct identification of the technological distance to others can aid considerably in their search for new technologies and collaborations.

4.2 Board Mobility

In recent years (and actualized in 2009) the Science and Research Based Business program (part of the Faculty of Science of the University of Leiden) conducted a study to determine the different work histories of all members of the Board of Directors of the firms present at the park at that time (Jousma and Van Rossum, 2009). Although this dataset does not give exact years of employment history, it does give insight into the organizations that the members worked for, generally in chronological order. In the study of De Groot (2011) this dataset is enriched with manually found data, and therefore becomes useful for labour mobility analysis in a similar manner as is explained in section 4.1.1. The work history on the park comprises an extensive set of firms, which not all are present in the data about technological relatedness (as not all firms applied for patents). As both datasets have to be regressed against one another, they need to cover the same firms. From the labour mobility of board members this means that only the mobility between firms that have applied for patents is selected as dependent variable for the regression (see paragraph 4.3). A similar approach as in 4.1.1 with co-occurrences of labour between firms is then applied to create a network of connected firms through labour mobility. The next paragraph explains the final regression method to test how technological relatedness can explain labour mobility.

4.3 Regression Methodology

The in the theoretical framework's suggested relationship between technological distance (the inverse of technological relatedness) and labour mobility can now be tested using the data created in the previous sections. In total 3 different regressions will be performed where in all cases the logarithm of the technological relatedness between firms is used as the independent variable, and the labour mobility of inventors, labour mobility of members of the Board, and both groups combined respectively will be used as the dependent variable in the regressions. The method applied for regressions can however not be the standard Ordinary Least Squares (Pindyck and Rubinfeld, 1998) regressions, as problems of autocorrelation are likely to persist. The University may for instance have a wider range of technologies in their patents than many other organizations on the park. This would make the technological distance between the University and the other actors significantly higher overall, and is known as autocorrelation (Pindyck and Rubinfeld, 1998). More formally: "...observations in network data have varying amounts of dependence on one another according to which row or column they 'belong'." (Krackhardt, 1988, p. 361)

Although autocorrelation is hypothesized, it is more difficult to test for dyadic data. Nevertheless a series of experiments by Krackhardt (1988) shows how the Qaudratic Assignment Procedure (QAP) is in almost every simple and multiple regression model the best option to avoid autocorrelation and superior to OLS in network analyses. In this procedure the matrix is scrambled in rows and columns, but both in the same fashion. This prevents scrambling of the technological distances of firm A keeping the dependence between the columns and rows remains intact. When several of these permutations are performed the standard error becomes independent from the observations providing unbiased results, and the resulting coefficient is in case of significance (p-value < 0,05) unlikely to originate from chance (Simpson, 2001). The correlation coefficients are computed in a similar way (synchronous permutation of the rows and columns), indicating how the two variables significantly overlap. Furthermore, applying this method in statistical software can be significantly more practical in case of large datasets, where other methods might require to tabulate each combination of the matrices of the dependent and independent variable.

To control for an 'over presence' of the University of Leiden and the Leiden University Medical Centre, additional dummy variables DumUni and DumLUMC are also included. They represent the same matrices as the dependent and independent variable, but are 1 for each combination with the University or the Medical Centre, and 0 otherwise. The significance of the coefficients shows whether the inclusion of these dummies was justified (they are discarded in case of insignificance). The network analysis software package UCInet (Borgatti et al., 2002) provides a tool for performing QAP-correlations and QAP-regressions. Whenever the right matrices as dependent and independent variable (matrices of the same shape with combinations of the same set of observations, in this case firms) are used as input, the program calculates different correlations and a regression output with a goodness-of-fit measure (R^2) and significance statistics. For the results see paragraph 5.3.

5. RESULTS

The results obtained from applying the methods described in chapter 4 will be described and discussed in the following two chapters. First, in this chapter the results are presented using the appropriate network graphs and statistics. Chapter 6 will then discuss the results by comparing them with the theoretical arguments that were made in the theoretical framework in chapter 2. This chapter is structured as follows: the first paragraph elaborates on technological relatedness using the patent data and the accompanying IPC-codes. Several descriptive statistics give insight into the features of the data, and by using the method of Engelsman and Van Raan (1991) several technology maps of the Leiden Bioscience Park are depicted. The second paragraph discusses the data of labour mobility, both inventor and Board mobility. The third and final paragraph of this chapter concerns the final QAP-regression, and with that answers the research question of this study.

5.1 TECHNOLOGICAL RELATEDNESS

The patent dataset that is created according to the criteria mentioned in chapter 4 comprises in total 681 unique patents. In the following section the descriptive statistics of these patents are discussed, including the organizations that applied for the patents and the technology codes that characterize the patents. In section 5.1.2 the results of the technological relatedness study are presented.

5.1.1 DESCRIPTIVE STATISTICS

Table 4 gives an overview of the number of patents that where applied for per year. Following the changes in the patent applications it provides a good estimation of the growth of the park over the years. The first 10 years after the establishment of the park in 1984 indicate some moments of significant growth, especially in 1988, but are relatively constant with respect to patent applications. However, from 1995 on the park experiences higher levels of patent applications until 2007, which is the last year of the regionalized REGPAT patent database. The patents of the remaining years are included manually as was described in paragraph 4.1.

From the regionalized patent database and a complementary manual search a list of applicants is subtracted of organizations that were ever located on the park since it was founded, and applied for at least one patent. This list can be found in table 15 in the appendix.

The larger share of patents on the park are applied for by a single applicant (86%), the other patents have more than one applicant. Half of those patents are owned by the University of Leiden. From these collaborations 17 patents are collaborations within the park, all with 2 applicants. These collaborations are not excluded from this table, this explains the total number of patents per applicant to be 17 patents higher than the patent count in table 15. The number of unique patents on the park thus remains to be 681. These collaborations will further be discussed in paragraph 5.2 together with the labour mobility.

Applicants with the largest number of patents are the University of Leiden, Crucell, Genencor International, Centocor Ortho Biotech and the Leiden University Medical Centre. An applicant has on average almost 18 patents, but due to the large number of small patent holders (26 applicants have less than 5 patents) and the few large patent holders as mentioned above the standard deviation is relatively large at 34,5.

Each of these patents comprise a certain invention in a technological field. These technologies are denoted by the International Patent Classifications, or IPC. In this study, the focus is on the first 7 digits, as is explained in paragraph On average, a patent is characterized by 4,86 IPC-codes, with a standard deviation of 4,19 (see table 16). The technology code that is designated to a patent the most times is C12N015 with a frequency of almost 600. Note however, that in this case a single patent can have a single IPC multiple times. This is because after discarding the different subgroups, the first 7 digits can be equal. In the study for technological relatedness these doubles have been discarded to avoid biasedness of the results. When these doubles are discarded the IPC-code C12N015 remains the most frequent one, but is halved in frequency. 300 of the 681 patents thus have this technology code, which makes it the most prominent technology of the park. This will also be shown in the network graphs of the next section (the top 25 is in table 17).

5.1.2 TECHNOLOGICAL RELATEDNESS

Following the methodology of Engelsman and Van Raan (1991) as is discussed in chapter 4 a certain relative distance between different technologies can be calculated. By measuring co-occurrences of IPC-codes in patents a 1-mode matrix can be formed with IPC-codes on both the horizontal and the vertical axis. This matrix is symmetrical in its diagonal, and each combination between two IPC-codes can be 0 whenever both IPC-codes do not co-occur in a patent, or 1 or larger when these IPC-codes do co-occur in one or more patents. From this matrix a network graph can be drawn, which essentially displays the technology field of the LBSP, and how related these technologies are. In this section a comparison of this technology field over the last 25 years will be made using the network graphs and measures of centralization of the network.

The patent data range from 1984 to 2010, with in the last 2 years only the patents that were obtained according to the manual search as is described in chapter 4. This timeframe is divided into 5 segments: 1984-1988, 1989-1993, 1994-1998, 1999-2003, and 2004-2010. The network graphs of the most prominent IPC-codes of each period are in figure 4-8, respectively. A network graph of the entire time span 1984-2010 and all IPC-codes is included as well as figure 10 in the appendix. Although the figure is very detailed and therefore difficult to read, it does show that the network is comprised of 4 parts: the main component, two smaller components which are unrelated to the main component, and the four IPC-codes in the upper-left corner which are not related to any other IPC-code.

Comparing the graphs of the time cohorts shows a similar trend as can be seen in table 16. As the number of patents increase, this also increases the number of IPC-codes, and therefore the size of the network. However, it also shows that the increase of patents not only increases the volume of the technologies on the LBSP, but also the diversity of the technologies. When over the years the technologies would remain the same the network graph would have a more similar shape over time, as the size of the co-occurrences are not incorporated in these graphs. This can be further explored using more comparable centralization-measures of the different networks.

Table 1: Top 10 IPC normalized degree centrality, per timeframe (micro level)

	1984-1988		1989-1993		1994-1998		1999-2003		2004-2010		1984-2010	
1	C12N015	0,62	C12N015	0,60	C12N015	0,53	C12N015	0,52	C12N015	0,26	C12N015	0,44
2	C12N001	0,48	C12N009	0,39	C07K014	0,33	C07K014	0,32	A61K039	0,24	C07K014	0,25
3	C12R001	0,33	C12N001	0,23	C12N009	0,26	C12N005	0,26	C07K016	0,20	C12N005	0,19
4	C12N009	0,29	A01H005	0,21	C12N005	0,19	A61K048	0,23	A61K038	0,19	A61K039	0,16
5	A01H001	0,24	C12R001	0,19	C12Q001	0,16	A61K039	0,18	G01N033	0,18	C12N009	0,16
6	C12N005	0,24	C07K014	0,16	A61K038	0,15	A61K038	0,16	C07K014	0,17	A61K038	0,16
7	C07K014	0,19	C12P021	0,16	C12N001	0,15	G01N033	0,14	C12N005	0,11	G01N033	0,15
8	A01H005	0,14	G01N033	0,16	A61K048	0,14	C12N009	0,11	C12Q001	0,11	C07K016	0,13
9	A01K067	0,14	C12N005	0,14	C07K016	0,12	C07K016	0,10	C12N009	0,09	A61K048	0,12
10	C12P021	0,14	A01K067	0,11	G01N033	0,11	C12Q001	0,10	A61K031	0,08	C12Q001	0,11

Degree centralization is applied on three levels, as is argued in section 4.1.2. The micro level is on the level of the individual IPC-code. Table 1 gives the top 10 IPC-codes for each timeframe (and the total timeframe) with the highest degree centrality. Changes within this top 10 are marked light and dark grey; the light grey highlights the IPC-codes which were also present in the top 10 of the previous timeframe , the darker grey the IPC-codes which are new in the top 10. The unmarked IPC-codes are present over the entire time span. The table shows that in the first 10 years of the LBSP the changes in the most prominent technologies were relatively small, while the next 10 years indicate how some other technologies have become significantly more prominent. In the timeframe 1999-2003 only 4 of the technologies that were in the top 10 at the start of the park remain in the top 10. Interestingly, the most prominent technology, C12N015, remains its top position throughout the entire time span. The figures 4 to 8 show these top 10's in network format.

Figure 4: IPC relatedness 1984-1988, top 10

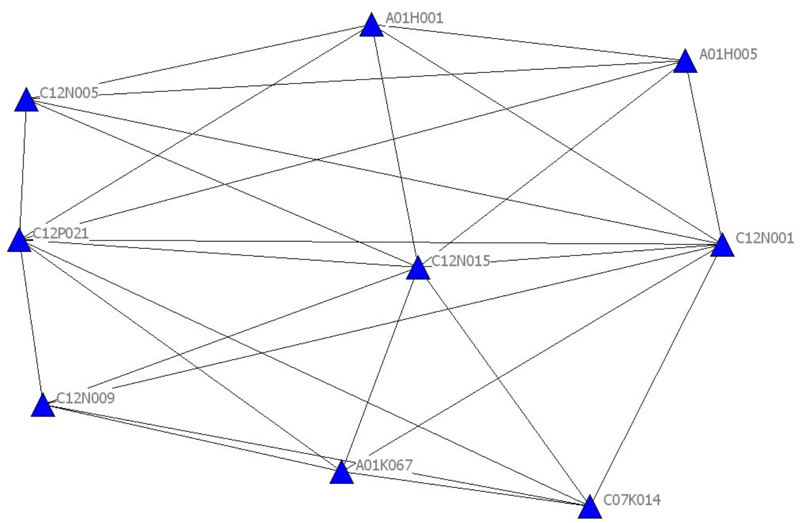


Figure 5: IPC relatedness 1989-1993, top 10

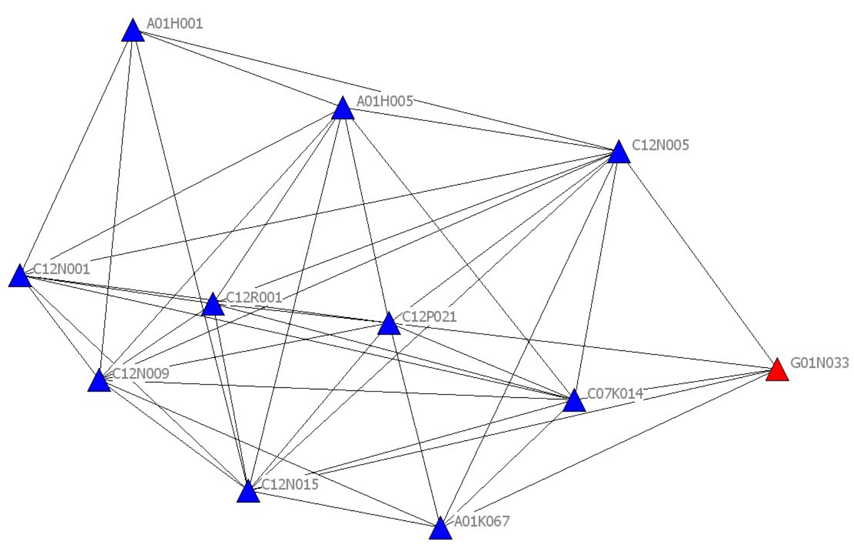


Figure 6: IPC relatedness 1994-1998, top 10

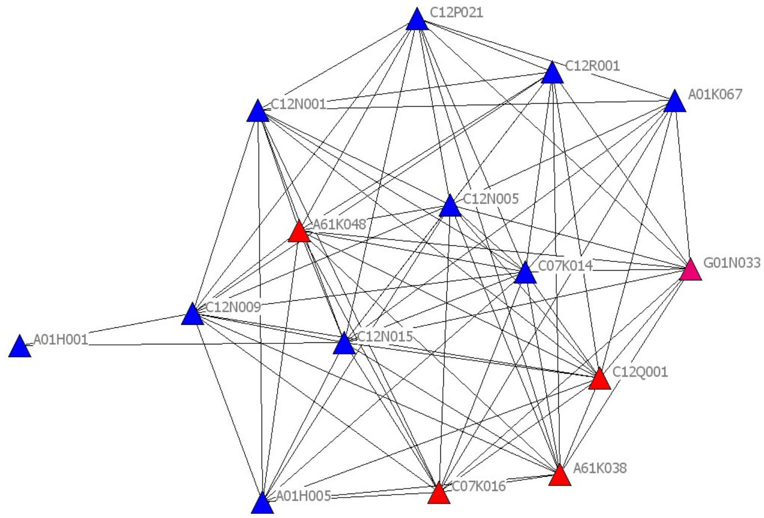


Figure 7: IPC relatedness 1999-2003, top 10

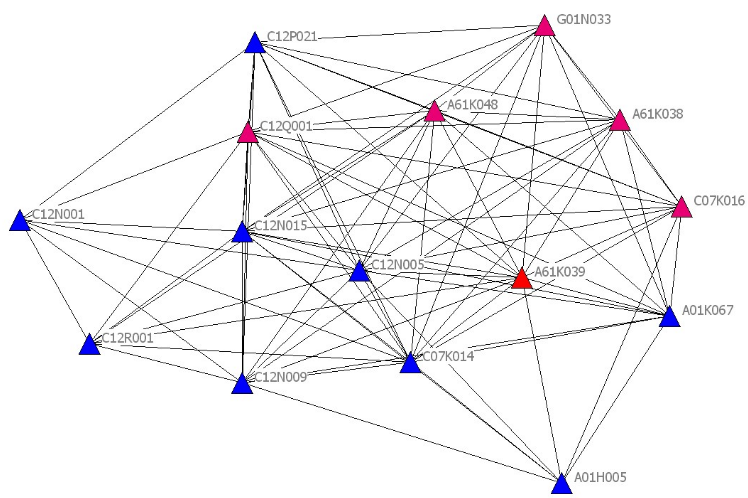
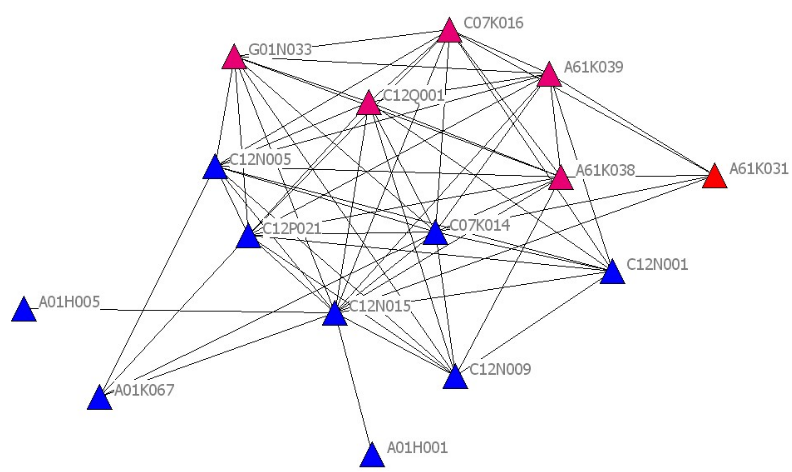


Figure 8: IPC relatedness 2004-2010, top 10



The blue nodes are the top 10 of the first cohort, the red ones are new in the top 10 (similar to the darker grey in the table) and the purple ones are also new but were already present in the previous cohort (similar to the light grey in the table). In the period 1999-2003 the IPC-code A01H001 disappeared from the LBSP entirely, but it appeared again in the last period. The code C12R001 disappears in the last period from the park. These changes illustrate the technological evolution of the park over time.

The meso-level of degree centralization in table 2 gives a broader overview of how the degree centrality changed over the years. The average centralization decreases over time, but less fast than the standard deviation. Although the individual centrality of the IPC-codes decreases on average, the differences between these centralities decrease with a slower rate and thus become relatively larger. The standard deviation has the relative largest size in the entire network over all years, but in absolute terms the smallest size (together with the average degree). Concentration of the network thus decreased over the years, but the relative differences among the degrees of the IPC-codes increased.

Table 2: Average normalized degree centralization, per timeframe (meso level)

	1984-1988	1989-1993	1994-1998	1999-2003	2004-2010	1984-2010
Average	0,1140	0,0623	0,0398	0,0295	0,0291	0,0151
Standard deviation	0,1289	0,0929	0,0761	0,0668	0,0527	0,0437
Difference	0,8846	0,6709	0,5229	0,4423	0,5528	0,3460

The macro-level gives the degree centralization of the entire network, taking into account both the number of IPC-codes and the number of patents, and therefore comparable over the different networks over time. Following the methodology of Borgatti and Everett (2005) as is proposed in section 4.1.2, the degree centralization of the networks is calculated and summarized in table 4.

Table 3: Degree centralization, per timeframe, in percentages

	1984-1988	1989-1993	1994-1998	1999-2003	2004-2010	2004-2007	1984-2010
Centrality	42,82	34,66	27,03	25,93	13,03	11,21	16,17

Centralization was highest at the start of the LBSP, and has decreased since then. The network was in the first 10 years not only smaller, but also more concentrated around a small number of technologies. The smaller core and larger distance to the periphery of the network changed over time to a larger core and a smaller distance to the periphery. Nevertheless, the low level of centralization in the final cohort is striking, especially as the overall centrality over the entire time span is larger. As hypothetically the incomplete data at the final years of this cohort could influence the results, the centrality of the years that were also included by REGPAT (2004-2007) is also calculated . The even lower centrality of 11,21 does not explain the first results, but might in fact even emphasize the implications of this lower centrality.

5.2 LABOUR MOBILITY

The mobility of labour on the LBSP is measured by using two complementary sources, where the first one measures inventor mobility and the second one mobility of members of the Board of Directors of the firms.

5.2.1 INVENTOR MOBILITY

The regionalized patent data of the OECD (REGPAT) provides for each regionalized patent the names and addresses of the inventors that worked on that patent. On average, 3,42 inventors have developed a patent, with a standard deviation of 2,18. Using the same selection criteria as is used for determining the IPC-codes in section 5.1.1, a list of inventors is created. This list comprises in total 957 unique inventors, predominantly Dutch (see table 4).

Table 4: Inventors per country

Country code	Inventors
AT	2
AU	4
BE	23
CA	3
CH	3
DE	25
DK	15
ES	1
FR	2
GB	34
IT	7
NL	606
NO	2
NZ	3
SE	4
US	223
Total	**957**

A third of the inventors is foreign. This has two major reasons: a co-author of a patent developed on the LBSP might be foreign or at least have a foreign nationality and residency, and patents are also included from firms located on the park, but where patents are only applied for through their headquarters abroad. As is explained in the methodology, it is assumed that for those patents belonging to the entire firm the ones with a Dutch inventor are likely to be (at least partially) developed on the park, indicating the presence of this knowledge and technologies in the division of the firm on the LBSP. The foreign co-authors of these patents are not excluded from the dataset. These 957 inventors developed on average on 2,34 patents. The top 10 of inventors with the most patents can be found in table 5. Two of the ten are located in the United States, the most prominent ones are from the Netherlands.

Table 5: Top 10 inventors with most patents

	Inventor	Patents
1	Bout, Abraham (NL)	35
2	Havenga, Menzo Jans Emco (NL)	32
3	Vogels, Ronald (NL)	27
4	Quax, Wilhelmus Johannes (NL)	19
5	Jones, Brian Edward (NL)	18
6	de Kruif, Cornelis Adriaan (NL)	16
7	Melief, Cornelis Johannes Maria (NL)	16
8	Giles-Komar, Jill (US)	15
9	Hooykaas, Paul Jan Jacob (NL)	15
10	Scallon, Bernard (US)	15

Mobility of inventors is assumed when an inventor applies for multiple patents at multiple applicants. Of the 957 inventors 81 (8,5%) have done this, all of them only moved once. To identify the applicant with the greatest centrality of the network the degree centrality of the applicants is computed by counting their connections with the other applicants. Based on the application years of the patents the direction of movement is determined.

In the degree centralization the connections between the applicants because of a collaboration is also included. These involve 17 connections, where 16 are between two different applicants and 1 is between two divisions of the same firm. The top 10 firms with the highest centrality degree (sorted left by OutDegree and right by InDegree) are to be found in table 6.

In total 19 of all 36 applicants are connected in some way to another applicant. An applicant has on average 2,79 in or outgoing connections (standard deviation OutDegree and InDegree is 7,70 and 6,97 respectively). The University of Leiden has the largest centrality in this network for both directions of inventor movement, even when its 11 collaborations with other applicants would be discarded.

Table 6: Top 10 applicants with highest centrality degree (through inventors and collaborations)

#	Applicant	OutDegree	InDegree	Applicant	#
1	University of Leiden	43	34	University of Leiden	1
2	Syngenta Mogen B.V.	17	23	Leiden University Medical Centre	2
3	Crucell Holland B.V.	13	13	Syngenta Mogen B.V.	3
4	Leiden University Medical Centre	11	13	Crucell Holland B.V.	4
5	TNO	6	6	TNO	5
6	OctoPlus Technologies B.V.	4	4	OctoPlus Technologies B.V.	6
7	Boston Clinics PDT B.V.	4	4	Prosensa B.V.	7
8	Pharming Group NV	3	4	Photobiochem N.V.	8
9	Galapagos Genomics B.V.	2	3	Flexgen Technologies B.V.	9
10	Prosensa B.V.	2	2	Pharming Group NV	10

5.2.2 Board mobility

The mobility of the members of the 2009's Board of Directors does not imply the same thing as the inventor mobility, but can (to a lower extent) still be a source of knowledge spillovers. Although the mobility network comprises 112 firms and the University, only firms that are also present in the REGPAT database can be used in this study (36, see section 4.1.3). A more in-depth study of this network is done by De Groot (2011).
After selecting the appropriate data, 18 members of the Boards have moved between organizations which are also present in table 15. The top 10 of degree centrality is provided in table 7.

Table 7: Top 10 applicants with highest centrality degree (through Board members)

#	Applicant	OutDegree	InDegree	Applicant	#
1	Crucell Holland B.V.	6	3	Prosensa B.V.	1
2	Pharming Group NV	2	3	Galapagos Genomics B.V.	2
3	OctoPlus Technologies B.V.	2	1	Crucell Holland B.V.	3
4	Syngenta Mogen B.V.	2	1	Pharming Group NV	4
5	Prosensa B.V.	1	1	OctoPlus Technologies B.V.	5

6	University of Leiden	1	1	University of Leiden	**6**
7	Biofocus DPI B.V.	1	1	Leiden University Medical Centre	**7**
8	Xendo Holding B.V.	1	1	Ingeny B.V.	**8**
9	Genencor International, Inc.	1	1	Flexgen Technologies B.V.	**9**
10	Galapagos Genomics B.V.	1	1	TNO	**10**

The role of the University is smaller than it is in table 6, while in table 7 Crucell takes a more prominent place. But as both the in and outdegrees are relatively small overall (on average 0,47), the implications of table 7 seem irrelevant.
The three proposed relations (including collaborations on a patent) are drawn as a network graph in figure 9. The figure shows that only 2 connection between applicants overlap for Board mobility and inventor mobility. The University of Leiden, the Leiden University Medical Centre, Crucell Holland B.V. and Prosensa B.V. have the most connections with different applicant, some more intense than others. However in the case of stronger links (more labour mobility between applicants) either the University of Leiden of the Leiden University Medical Centre is involved. This figure highlights their central position in the Leiden Bioscience Park.

5.3 QAP REGRESSIONS

The technological relatedness based on co-occurrences is the main source of measuring technological relatedness of firms. The geodesic distance between the technologies is therefore 'lifted' via the patents to the firm level, see the methodology in section 4.1.3.
The result is a firm by firm matrix, each of them connected by a certain technological distance. The larger this distance, the more unrelated firms are technologically. On average, the distance between two firms is 5,7, the standard deviation is 12,1. The technological distance thus has a wide range, from a minimum of 0,58 to a maximal distance of 114,39. Taking the natural logarithm to enable log-linear interpretation of the results decreases these absolute differences.
Several separate QAP's will be performed. In all tests the logarithm of the technological distance is used as the independent variable with the inventor mobility, Board mobility and the summed mobility as the dependent variable. In these tests the dummy variables for the University and the LUMC is included, a test of the summed mobility and the technological distance without using the dummies is also performed to illustrate the impact of both applicants.

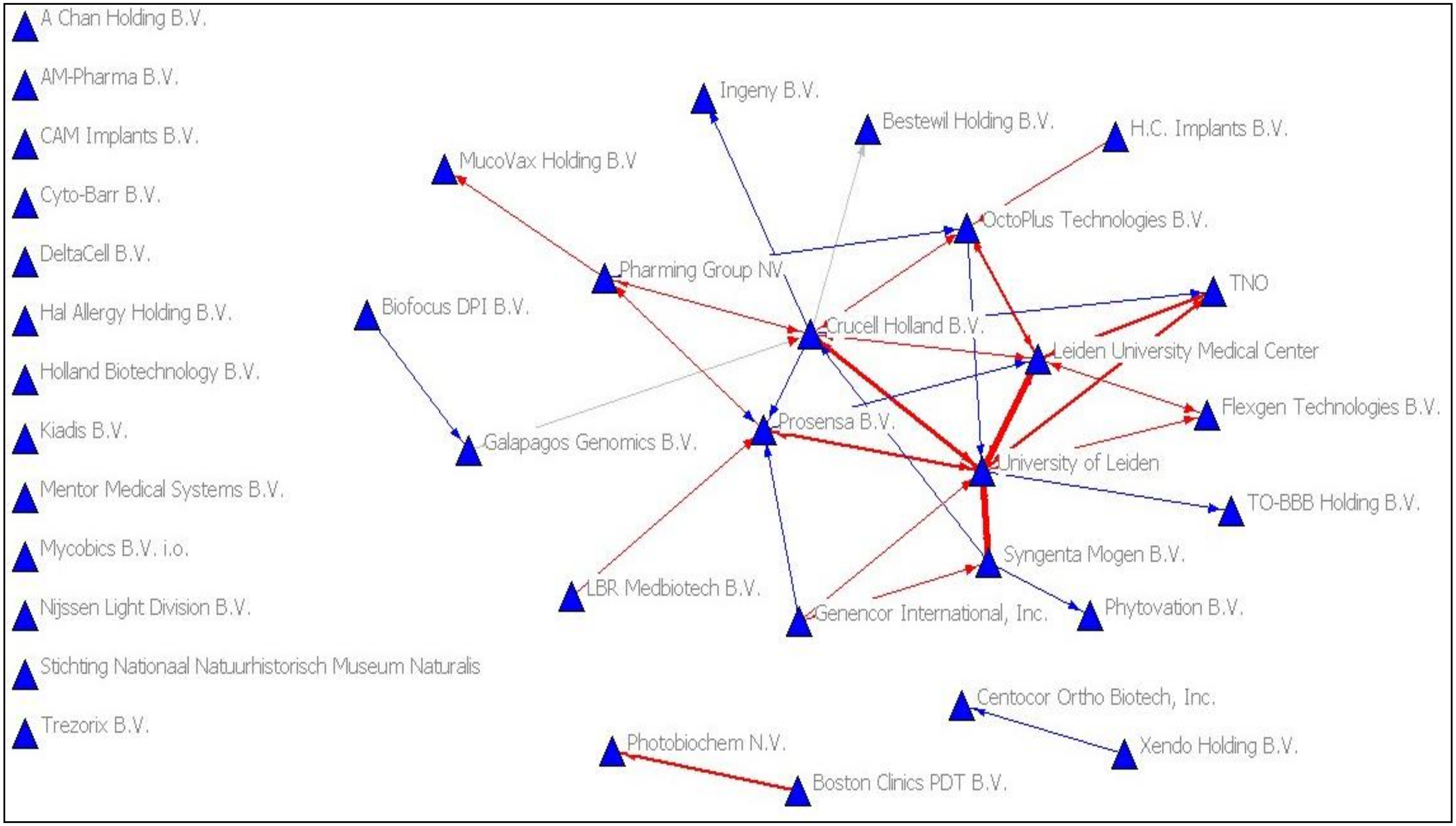


Figure 9: Labour mobility

Marginal note:
Blue is Board mobility
Red is inventor mobility
Grey is both inventor and Board mobility (overlap)
Thickness of line is relative size of mobility

Of each test the correlation coefficient, the R^2 and the beta is computed. The results are summarized in table 8 and 9.

Table 8: QAP correlations variables

	Appl TechDist	Board Mob	Inventor Mob	Total Labour Mob
Applicant TechDist	1,00	-0,005	0,19**	0,19**
Board mobility	-0,005	1,00	0,04*	0,17**
Inventor mobility	0,19**	0,04*	1,00	0,99***
Total Labour Mobility	0,19**	0,17**	0,99***	1,00

Note: * is significance at 10%, ** is significance at 5%, * is significance at 1%**

The correlation matrix in table 8 indicates both significant positive correlation between inventor mobility and the total of labour mobility with the dependent variable, Board mobility has an insignificant correlation of almost zero. The Board mobility itself does not seem to be in line with the technological distance between the applicants, but as the extent of Board mobility is relatively low in comparison with inventor mobility (18 of 113 movements is because of Board mobility) it has relative little effect on total labour mobility's correlation with the technological distance. This can also be observed in the low correlation of Board mobility with total labour mobility (0,17).
In the final QAP regression the main research question is answered. The regressions tests whether the technological distance between the applicants has an effect on the labour mobility between those applicants, and uses the two dummies to check for biased results caused by the University or LUMC. The results are summarized in table 9. The technological distance has a negative effect on the mobility of the Board members, although the coefficient is small and insignificant. The R-squared shows that the variable does not explain the model significantly, as it is very close to zero. The effect of technological distance on inventor mobility and total labour mobility is almost equal; the coefficients, significance and goodness of fit are close to identical. Their results are therefore discussed simultaneously.
The R-squared of 0,125 of inventor mobility means that 12,5% of the dependent variable is explained by the independent variable. Interestingly the R-squared is slightly smaller for the larger dataset of labour mobility, which means that Board mobility does not help in explaining the dependent variable.
The dummy variables are significant and positive, indicating that both the University and the LUMC significantly affect the results. Nevertheless, the coefficients of Log(TechDist) are also positive and

significant. The significance of the other applicants in the regression is therefore not jeopardized by either the University or the LUMC; the results still hold when they would have been excluded completely.

The positive coefficient implies that as the technological distance increases, labour mobility also increases. The fact that for technological distance the logarithm is used allows a more accurate interpretation of the coefficient: when technological distance increases by 1%, the number of linkages between applicants in the form of labour mobility increases by 0,30. Although the relatively low R-squared implies that other forces are at work in explaining labour mobility, the coefficient is significant and substantial. As this is not in line with the expectations formulated in chapter 2, these results create room for discussion and may have implications for economic theory concerning technological relatedness and knowledge spillovers.

Table 9: QAP regression with dummies

	Model 1: Board mobility	**Model 2: Inventor mobility**	**Model 3: Total Labour mobility**
Intercept	0,0284	-0,1251	-0,0967
Log(TechDist)	-0,007971 (0,017131)	0,305683* (0,753218)	0,297713** (0,713515)
DumUni	0,035625 (0,021554)	1,691709** (0,401233)	1,727334** (0,295039)
DumLUMC	0,00551 (0,029899)	0,718666** (0,663447)	0,724176** (0,699308)
R^2	0,002	0,125	0,124
No. Observations	**1260**	**1260**	**1260**

Note: * is significant at 1%, ** is significant at 5%, standard deviations in parentheses.

To illustrate the effect of the dummy-variables on the results the regression is also performed without the dummies. The results can be found in table 18 in the appendix. The coefficients between the technological distance and the inventor mobility and the total labour mobility are more than twice as large compared to the results with the dummies, for the Board mobility the results do not change much. The R^2's however are significantly lower, indicating a poorer goodness-of-fit. Using the dummies as is demonstrated above thus gives a better view of the dynamics on the LBSP.

6. CONCLUSION AND DISCUSSION

The results obtained from the empirical study in chapter 5 give an unambiguous and direct answer to the research question of chapter 2. The research question is: *how is labour mobility between two firms affected by their technological distance?* The results show that the technological distance positively affects the labour mobility, indicating that organizations with a larger technological distance share more labour than organizations that are technologically closer to each other.

In this chapter this conclusion and the other results obtained in the previous chapter are discussed and linked to the different theories formulated in chapter 2. The results are discussed on three different levels: the level of the park, of the firm and of the individual (either inventor or member of the Board of Directors).

The analysis of the development over time of the park indicates that the park had only little growth in the first 10 years after its establishment, but grew significantly faster from then on. The nature of the growth based on the technology codes of the patents is a continuous diverging of technologies. Rather than applying more of the same technologies, more new technologies are introduced over the years. Especially in the last 6 years the centralization decreased dramatically, meaning a larger distance between the core-technologies on the park and the other, peripheral technologies. By expanding the portfolio of technologies the park avoids a lock-in as formulated by Boschma (2005) and Frenken (2007). The park is less dependent on a small number of technologies, and cross-fertilization can be achieved by increasing the diversity. The continuous inflow of knowledge from other countries through settlement of foreign firms and inventors contributes to the diversity. However, almost all technologies are related at some extent (see figure 9), indicating that the diversity remains close to the core-business of the park.

On the level of the firm or organization a similar prevention of lock-in seems to exist. The regression in paragraph 5.3 shows how technologically related firms share less inventors than firms that are relatively unrelated. As inventors are an important source of knowledge for a firm, firms will not randomly hire inventors; instead they will look for inventors possessing the knowledge that benefits the firm the most (Breschi and Lissoni, 2001b). The results indicate that firms benefit more from relatively less related knowledge. However, all actors on the park belong to a network of related technologies. This suggests a common science base (Boschma, 2005; Feldman and Audretsch, 1999) on the park, where firms rather look for complementary than substitutionary knowledge. The modified inverted U-shape suggested in paragraph 2.1 seems to reflect the results the most, although labour mobility would fit the vertical axis best. A low level of technological variety is not optimal as it could cause a lock-in, and a too high level of variety would diminish the advantages of the common science base. Nevertheless, it seems that a threshold of relatedness has been passed, and the focus of the park is no longer specialization but diversification within the common science base. As the park is continuously expanding and diversifying over the last 10 years it seems that the optimal level of technological variety or diversification has not been achieved (yet), and the present point of variety is above the threshold of the lock-in, but still below the optimal level.

The level of the individual has not been the centre of this study so far, but the regression results have interesting implications. The results for the mobility of the members of the Board of Directors is insignificant, implying that here the technological relatedness does not play a role. As a Board member generally does not carry detailed technological knowledge but is more relevant to a firm because of his management skills, the technologies of the previous firm of this person is of less relevance than for an inventor. The mobility of the inventor is according to the regressions results in most cases between relatively unrelated firms or organizations. Next to the diversification interest of the firm as described above, the inventor has an influence in its mobility as well. Assuming it will pursue a maximum economic rent, it prefers a firm where its knowledge is not applied yet so it has monopoly power over this knowledge in this firm. Further studies into the incentives of the inventor and its utility might confirm this suggestion.

The reversed causality of section 2.1.4 where hypothetically an inventor might also have an effect on the technological distance between firms as it enables convergence of their knowledge bases is already theoretically marginalized, the positive coefficient emphasizes this. If the reversed causality would play a role in practice, this would imply a negative effect on the coefficient as it is in the regression (more mobility would decrease the technological distance between two firms), making the effect of technological relatedness on labour mobility even stronger to compensate for the (hypothetical) negative effect of the reversed causality. Further study of this relationship can test what the actual effect is of labour mobility on the technologies of the firms. Other limitations and recommendations are described in the next section.

LIMITATIONS AND RECOMMENDATIONS

The results of this study show a significant and interesting relationship, but at the same time poses more questions and lines for further research. The most central actor in the network of labour mobility and the most distant actor in the network of technologies is the University. Although the regression controlled for its impact by using dummy variables, its role in the park is not the central subject of this study. Nevertheless, its centrality in the labour network emphasizes the importance of the University for inventors and for other organizations collaborating with the University in the application of patents. The fact that it is also the technologically most distant actor with respect to the other firms and organizations on the park lies in the nature of University research and the measurement of co-occurrences of IPC-codes. As a University has by definition a broader research base it applies more different technologies (and more IPC-codes) than a typical private actor will. This creates a larger technological distance with respect to other organizations. Furthermore, the assumption that co-occurrences of IPC-codes in patents is a measure of technological distance might not be the only thing co-occurrences imply. An IPC-code co-occurring in a patent might also be an indication of the phase of the particular production process that applies to the patent. As a University focuses more on fundamental knowledge that can be applied into a broad range of technologies and a firm concentrates more on applied, downstream knowledge, this might explain the relative large distance between the two actors. Further research with respect to the product chain on the LBSP combined with co-occurrences of IPC-codes in patents will shed more light on this discussion. Disentangling the University into different departments and faculties will result in more robust results.

The role of the firms that are new to the park can be studied in more detail. Their contribution to the knowledge base of the LBSP will be different than the incumbents, and might even be the major source of diversification, as is suggested by Orsenigo et al. (2001). Future studies can test where diversification of the technologies comes from, and whether incumbents are able to absorb the new technologies on the park while new entries facilitate the observed overall degree of technological de-centralization.

Other complementary study might also use control variables such as firm size and firm performance to exclude its impact on the results. Although the technological distance is weighted for number of patents and IPC-codes per patent, controlling for firm size over a time span of 25 years was beyond the scope of this research. A further detailing of labour mobility on the park will increase reliability of the results as well. As inventor mobility is entirely based on patent data, it is therefore likely that some mobility is unobserved. More mobility tracking on an individual level by using LinkedIn and other social network sources will enrich the data and produce more credible results.

Finally, the method applied in determining the technological distance between two or more firms can be used in more managerial applications. Creating an overview of technological distances to firms for potential collaborations can serve as a complementary tool in determining the right partner, especially when technologies are important to the industry or the collaboration.

ACKNOWLEDGEMENTS

This study could not have been made possible without the inspiration, contribution and comments of some people. First and foremost I thank my supervisor Sandra Phlippen for her efforts in setting up the workgroup and her comments and guidance during the entire process. I also thank Harmen Jousma in providing insights into the dynamics of the park, Anet Weterings for her help in retrieving the OECD patentdata, and Jan Verheijen of TNO for his kind help in determining the TNO patents that are created on the Leiden Bioscience Park. Finally I also thank (in alphabetic order) Edgar Louwerse, Fleur de Groot, Jelte Dijkstra, Leon Wiendels, Pieter-Bart Visscher and Thijs Wolters for their comments and suggestions during the group meetings.

APPENDIX

Table 10: Number of entries and exits on the LBSP

Period	Companies								Other		Total
	Type of entry (% of total entries)					total entries	total exits	total	Not-for-profit	education & research	
	start-up	spin-off	merger	division start	re-location						
- 1984						0	0	0	1	2	3
1984-1988	2 (33)	2 (33)	0 (0)	2 (33)	0 (0)	6	0	6	1	0	7
1989-1993	6 (40)	4 (27)	0 (0)	4 (27)	1 (7)	15	1	14	1	0	15
1994-1998	7 (54)	4 (31)	0 (0)	1 (8)	1 (8)	13	4	9	0	1	10
1999-2003	12 (40)	14 (47)	1 (3)	2 (7)	1 (3)	30	6	24	2	1	27
2004-2008	8 (27)	10 (33)	0 (0)	6 (20)	6 (20)	30	8	22	2	1	25
Total '84-'08	35 (37)	34 (36)	1 (1)	15 (16)	9 (10)	94	19	75	7	5	87

Source: Jousma et al. (2009)

Table 11: Number of people employed on the LBSP

Year	In companies						In other		Total
	In type of entry (% of total in companies)					Total in companies (% of total)	Not-for-profit	Education & Research	
	start-up	spin-off	merger	division start	relocation				
1985	0 (0)	21 (19)	0 (0)	88 (81)	0 (0)	109 (2)	136 (3)	4863 (95)	5108
1990	40 (10)	73 (17)	0 (0)	265 (63)	40 (10)	418 (7)	169 (3)	5080 (90)	5667
1995	39 (4)	157 (16)	0 (0)	309 (31)	495 (50)	1000 (15)	262 (4)	5348 (81)	6610
2000	76 (4)	228 (12)	86 (4)	858 (45)	668 (35)	1916 (23)	184 (2)	6347 (75)	8447
2005	175 (7)	206 (8)	282 (10)	1424 (53)	602 (22)	2689 (27)	121 (1)	7126 (72)	9936

Source: Jousma et al. (2009)

Table 12: Patent applications filed to the EPO

EP_APPLT_REG *(2 207 204 rows)*	
Appln_nr	Patent application number
Appln_id[1]	**PATSTAT Application id (September 2009)**
Publn_nr	Patent publication number
Person_id[1]	PATSTAT person identifier (September 2009)
Applt_name	Applicant's name
Address	Address
Reg_code	NUTS3 region code
Ctry_code	Country
Reg_share[2]	Share ≤ 1
Applt_share[3]	Applicant share ≤ 1
Reg_type[4]	Regionalisation method

EP_INVT_REG *(5 113 927 rows)*	
Appln_nr	Patent application number
Appln_id[1]	**PATSTAT Application id (September 2009)**
Publn_nr	Patent publication number
Person_id[1]	PATSTAT person identifier (September 2009)
Invt_name	Inventor's name
Address	Address
Reg_code	NUTS3 region code
Ctry_code	Country
Reg_share[2]	Share ≤ 1
Invt_share[3]	Inventor share ≤ 1
Reg_type[4]	Regionalisation method

EP_PRIO_IPC *(7 616 763 rows)*	
Appln_id[1]	**PATSTAT Application id (September 2009)**
Prio_year	Priority year (first filing)
App_year	Filing year
IPC[5]	List of IPC classes (8^{th} edition)

Source: OECD, Regpat database, January 2010

Table 13: 25 unconnected patents

Application ID	Applicant
16992704	University of Leiden
17124979	University of Leiden
17351954	University of Leiden
17354647	University of Leiden
17412971	University of Leiden
16024945	University of Leiden
16066302	University of Leiden
16544738	University of Leiden
17837899	Stichting Nationaal Natuurhistorisch Museum Naturalis
17468016	Nijssen Light Division B.V.
15954692	Dutch Space B.V.
15955419	Dutch Space B.V.
16225085	Dutch Space B.V.
16295778	Dutch Space B.V.
17668769	Dutch Space B.V.
17697209	Dutch Space B.V.
17746696	Dutch Space B.V.
17469363	Dutch Space B.V.
17470053	Dutch Space B.V.
17527629	Dutch Space B.V.
17668768	Dutch Space B.V.
16692853	OctoPlus Sciences B.V.
16480162	Produvation BV
17900896	Genencor International, Inc.
17906180	Genencor International, Inc.

Table 14: Patents per applicant, number of applicants per patent

Applicant name/Number of applicants	1	2	3	4	6	Patents/applicant
A Chan Holding B.V.	1					1
AM-Pharma B.V.	2					2
Bestewil Holding B.V.	3					3
Biofocus DPI B.V.	3					3
Boston Clinics PDT B.V.	2					2
CAM Implants B.V.	1	3				4
Centocor Ortho Biotech, Inc.	72	3		1	1	77
Crucell Holland B.V.	117	12	2			131
Cyto-Barr B.V.	1					1
DeltaCell B.V.	2					2
Dutch Space B.V.	11					11
Flexgen Technologies B.V.	1					1
Galapagos Genomics B.V.	3	1				4
Genencor International, Inc.	74	13				87
H.C. Implants B.V.	2					2
Hal Allergy Holding B.V.	1					1
Holland Biotechnology B.V.	2	3				5
Ingeny B.V.	3					3
Kiadis B.V.	8					8
LBR Medbiotech B.V.	1					1
Leiden/Amsterdam Centre for Drug Research LACDR	1					1
Leiden University Medical Centre	52	9				61
Mentor Medical Systems B.V.	1					1
MucoVax Holding B.V	1					1
Mycobics B.V. i.o.		1				1
Nijssen Light Division B.V.	2					2
OctoPlus Technologies B.V.	22	3	1			26
Pharming Group NV	25	3	1			29
Photobiochem N.V.	2	1				3
Phytovation B.V.	3					3
Produvation BV	1					1
Prosensa B.V.	4	1				5
Stichting Nationaal Natuurhistorisch Museum Naturalis	1					1
Syngenta Mogen B.V.	24	8	1			33
TNO	34	4				38
TO-BBB Holding B.V.	2					2

Trezorix B.V.	1					1
University of Leiden	63	67	5	3		138
Xendo Holding B.V.	2					2
Total number of patents	**551**	**133**	**10**	**4**	**1**	**698**
Average number of patents/applicant						**17,9**
Standard deviation						**34,5**

Table 15: IPCs per patent

Average	4,86
Standard deviation	4,19

Table 16: Patents per year

Year	Patents
1982	1
1983	1
1984	3
1985	4
1986	3
1987	2
1988	9
1989	11
1990	11
1991	11
1992	9
1993	15
1994	9
1995	28
1996	29
1997	28
1998	35
1999	46
2000	51
2001	58
2002	42
2003	37
2004	47
2005	59
2006	58
2007	54
2008	8

2009	9
2010	3
Total	**681**

Table 17: Top 25 IPCs

	IPC-7 digits	Frequency
1	C12N015	299
2	C07K014	171
3	C12N005	126
4	A61K039	110
5	C12N009	109
6	A61K038	108
7	G01N033	100
8	C07K016	88
9	A61K048	83
10	C12Q001	77
11	C12N001	63
12	A61K031	47
13	C12P021	44
14	A61P031	33
15	A01H005	32
16	C12R001	32
17	A61K035	31
18	A61P035	31
19	A61K009	30
20	C12N007	27
21	A01K067	24
22	C11D003	23
23	A61K047	22
24	A61P037	21
25	A61P009	19
	Total	**2307**

Figure 10: IPC relatedness all years

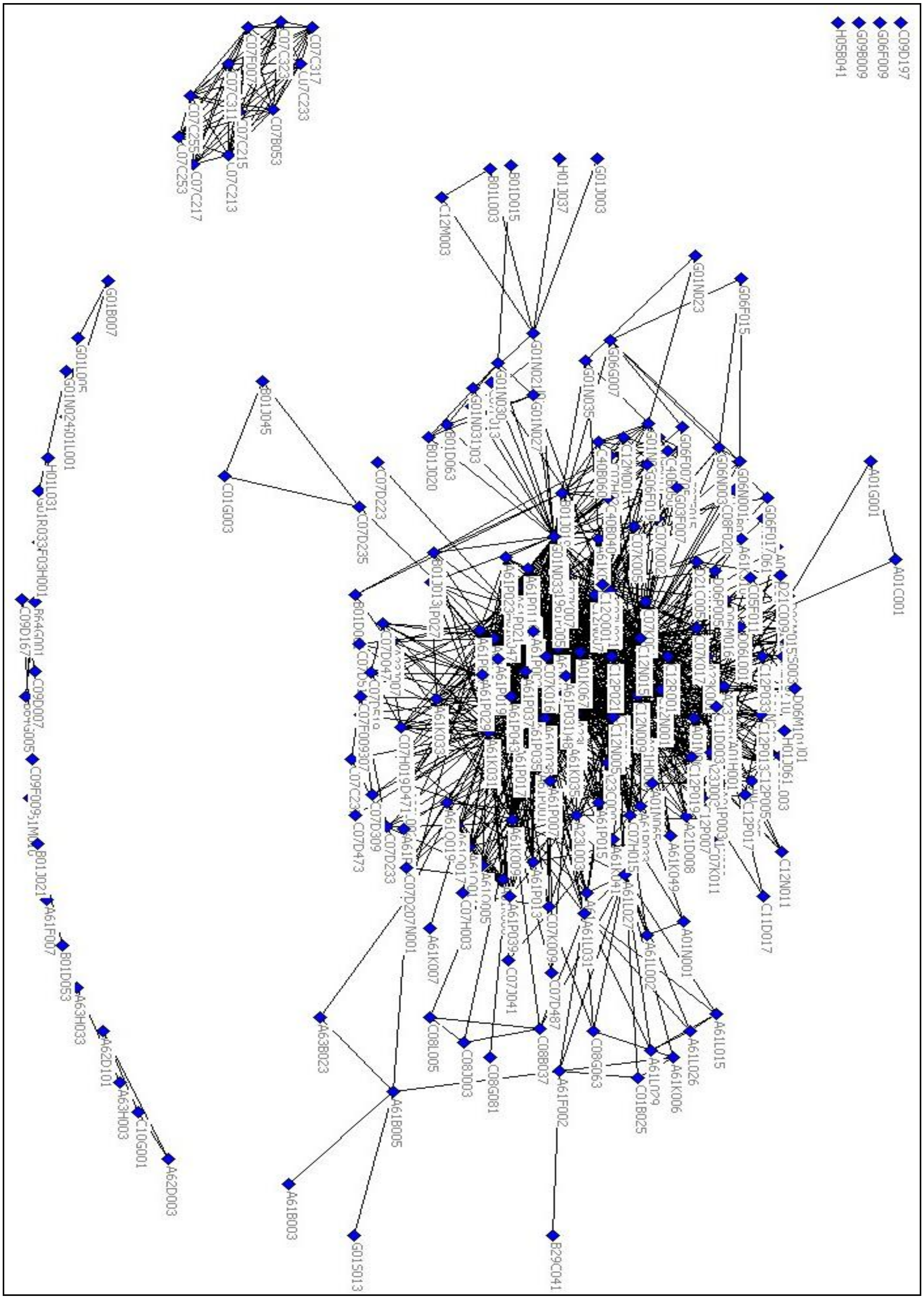


Table 18: QAP regressions without dummies

	Model 1: Board mobility	Model 2: Inventor mobility	Model 3: Total Labour mobility
Intercept	0,027998	-0,143341	-0,115343
Log(TechDist)	-0,002182 (0,013196)	0,633166* (0,141536)	0,630983* (0,062690)
R^2	0,000	0,036	0,035
No. Observations	**1260**	**1260**	**1260**

Note: * is significant at 1%, ** is significant at 5%, standard deviations in parentheses.

References

Articles:

Agrawal, A., Cockburn, I. and McHale, J. (2006). Gone but not forgotten: knowledge flows, labour mobility, and enduring social relationships. *Journal of Economic Geography,* **6**, pp. 571-591

Almeida, P. and Kogut, B. (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science,* **45** (7), pp. 905-917

Arita, T. and McCann, P. (2000). Industrial alliances and firm location behaviour: some evidence from the US semiconductor industry, *Applied Economics,* **32** (11), pp. 1391-1403

Arrow, K.J. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies,* **29** (3), pp. 155-173

Autant-Bernard, C. (2001). The geography of knowledge spillovers and technological proximity, *Economics of Innovation and New Technology,* **10**, pp. 237-254

Balsvik, R. (2006). Is mobility of labour a channel for spillovers from multinationals to local domestic firms? *Norwegian School of Economics and Business Administration,* discussion paper no. 25/06

Bathelt, H., Malmberg, A. and Maskell, P. (2004). Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation, *Progress in Human Geography*, **28** (1), pp. 31-56

Beaudry, C. and Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate, *Research Policy,* **38**, pp. 318-337

Boschma, R. (2004). Competitiveness of Regions from an Evolutionary Perspective, *Regional Studies,* **38** (9), pp. 1001-1014

Boschma, R. (2005). Proximity and Innovation: A Critical Assesment, *Regional Studies*, **39** (1), pp. 61-74

Breschi, S. and Lissoni, F. (2001a). Localised knowledge spillovers vs. innovative milieux: Knowledge "tacitness" reconsidered, *Papers in Regional Science*, **40**, pp. 255-273

Breschi, S. and Lissoni, F. (2001b). Knowledge Spillovers and Local Innovation Systems: A Critical Survey, *Industrial and Corporate Change*, **10** (4), pp. 975-1005

Breschi, S. and Lissoni, F. (2006). Mobility of inventors and the geography of knowledge spillovers. New evidence on US data, *CESPRI Working Paper 184*, Milan, Italy: Bocconi University.

Breschi, S., Lissoni, F. and Malerba, F. (2003). Knowledge-relatedness in firm technological diversification, *Research Policy*, **32**, pp. 69-87

Cohen, W.M. and Klepper, S. (1992). The Tradeoff Between Firm Size and Diversity in the Pursuit of Technological Progress, *Small Business Economics,* **4**, pp. 1-14

Döring, T. and Schnellenbach, J. (2006). What Do We Know about Geographical Knowledge Spillovers and Regional Growth? A Survey of the Literature, *Regional Studies*, **40** (3), pp. 375-395

Etzkowitz, H. (2002). Incubation of incubators: innovation as a triple helix of university-industry-government networks, *Science and Public Policy,* **20** (2), pp. 115-128

Feldman, M.P. and Audretsch, D.B. (1999). Innovation in cities: Science-based diversity, specialization and localized competition, *European Economic Review,* **43**, pp. 409-429

Frenken, K. (2007) Related variety, unrelated variety and regional economic growth, *Regional Studies*, **41**(05), pp. 685-697

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A. and Schleifer, A. (1992). Growth in Cities, *The Journal of Political Economy*, **100** (6), pp. 1126-1152

Griliches, Z. (1979). Issues in Assessing the Contribution of Research and Development to Productivity Growth, *The Bell Journal of Economics,* **10** (1), pp. 92-116

Guiliani, E. (2003). *Knowledge in the Air and its Uneven Distribution: A story of a Chilean Wine Cluster.* Presented at the DRUID Winter Conference 2003, Aalborg: Denmark

De Groot, F. (2011). *Social network formation and high tech clusters: two-way job mobility network analysis of the Leiden Bioscience Park*, Master Graduation Thesis, Erasmus School of Economics: Rotterdam, The Netherlands

De Groot, H.L.F., Poot, J. and Smit, M.J. (2008). Agglomeration externalities, innovation and regional growth: Theoretical perspectives and meta-analysis, *Department of Economics Working Paper Series, Number 1/08*, Hamilton, New Zealand: University of Waikato.

Von Hippel, E. (1994). “Sticky information” and the Locus of Problem Solving: Implications for Innovation, *Management Science,* **40** (4), pp. 429-439

Ibrahim, S.E., Fallah, M.H. and Reilly, R.R. (2009). Localized sources of knowledge and the effect of knowledge spillovers: an empirical study of inventors in the telecommunications industry, *Journal of Economic Geography,* **9**, pp. 405-431

Jaffe, A.B. (1989). Real Effects of Academic Research, *The American Economic Review*, **79** (5), pp. 957-970

Jaffe, A.B., Trajtenberg M. and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, *The Quarterly Journal of Economics*, **108** (3), pp. 577-598

Jousma, H., Scholten, V. and Van Rossum, P. (2009). Framework For Analyzing the Growth of University Research Parks Applied to the Bioscience Park in Leiden, the Netherlands, *Triple Helix VII International Conference 17-19 June 2009*, Glasgow: Scotland.

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

Krackhardt, D. (1988). Predicting with networks: Nonparametric multiple regression analysis of dyadic data, *Social* Networks, **10**, pp.359-381

McCann, P. and Simonen, J. (2005). Innovation, knowledge spillovers and local labour markets, *Papers in Regional Science,* **84** (3), pp. 465-485

Møen, J. (2005). Is Mobility of Technical Personnel a Source of R&D Spillovers? *Journal of Labor Economics,* **23** (1), pp. 81-114

Neffke, F. (2009). *Productive Places. The Influence of Technological Change and Relatedness on Agglomeration Externalities*. PhD thesis, Utrecht University: Utrecht, The Netherlands.

Orsenigo, L., Pammolli, F. and Riccaboni, M. (2001). Technological change and network dynamics. Lessons from the pharmaceutical industry, *Research Policy*, **30**, pp. 485-508

Owen-Smith, J. and Powel, P.W. (2004). Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community, *Organization Science*, **15**(1), pp. 5-21

Paci, R. and Usai, S. (1999a). Externalities, knowledge spillovers and the spatial distribution of innovation, *GeoJournal*, **49**, pp. 381-390

Paci, R. and Usai, S. (1999b). The Role of Specialization and Diversity Externalities in the Agglomeration of Innovative Activities, *Rivista Italiana degli Economisti*, **2**, pp. 237-268

Palomeras, N. and Melero, E. (2010). Markets for Inventors: Learning-by-Hiring as a Driver of Mobility, *Management Science*, **56** (5), pp. 881-895

Van Der Panne, G. (2004). Agglomeration externalities: Marshall vs. Jacobs, *Journal of Evolutionary Economics*, **14**, pp. 593-604

Van Der Panne, G. and Van Beers, C. (2006). On the Marshall-Jacobs controversy: it takes two to tango, *Industrial and Corporate Change*, **15** (5), pp. 877-890

Romer, P. M. (1986). Increasing returns and long-run growth, *Journal of Political Economy*, **94**, pp. 1002–1037

Shefer, D. and Frenkel, A. (1998). Local milieu and innovations: some empirical results, *The Annals of Regional Science,* **32**, pp. 185-200

Simpson, W. (2001). *QAP: the quadratic assignment procedure*. Paper presented at the *North American Stata Users' Group Meeting*. Available at: http://fmwww.bc.edu/RePEc/nasug2001/simpson.pdf (Accessed November 1st 2010)

Solow, R.M. (1957). Technical Change and the Aggregate Production Function, *The Review of Economics and Statistics,* **39** (3), pp. 312-320

Vacaro, A., Veloso, F. and Brusoni, S. (2009). The impact of virtual technologies on knowledge-based processes: An empirical study, *Research Policy,* **38**, pp. 1278-1287

Wilson, L. and Spoehr, J. (2010). Labour Relations and the Transfer of Knowledge in Industrial Clusters: Why do Skilled Workers Share Knowledge with Colleagues in Other Firms? *Geographical Research,* **48** (1), pp. 42-51

Zucker, L.G., Darby, M.R. and Armstrong, J. (1998). Geographically Localized Knowledge: Spillovers or Markets? *Economic Inquiry,* **36**, pp. 65-86

Books:

Audretsch, D.B. (2003). Globalization, Innovation and the Strategic Management of Places, in Bröcker, J., Dohse, D. and Soltwedel, R. (Eds.) *Innovation Clusters and Interregional Competition,* Springer-Verlag, Berlin, pp. 11-27

Audretsch, D.B. and Feldman, M.P. (2003). Knowledge Spillovers and the Geography of Innovation, in Henderson, J.V. and Thisse, J. (Eds.) *Handbook of Urban and Regional Economics - Volume 4,* Elsevier, Amsterdam, pp. 2713-2735

Borgatti, S.P. and Everett, M.G. (2005). Extending Centrality, in Carrington, P.J., Scott, J. and Wasserman, S. (Eds.) *Models and Methods in Social Network Analysis,* Cambridge University Press, New York, pp. 57-76

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

Krugman, P. (1991). *Economic Geography and Trade*. MIT Press: Cambridge, MA.

Marshall, A. (1920). *Principles* of *Economics, 8th edition*. Macmillan: London, United Kingdom.

Pindyck, R.S. and Rubinfeld, D.L. (1998). *Econometric Models and Economic Forecasts, Fourth Edition.* McGraw-Hill: New York, NY.

Polanyi, M. (1966). *The Tacit Dimension.* The University of Chicago Press: Chicago, IL.

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

Wasserman, S. and Faust, K. (1994). *Social Network Analysis*. Cambridge University Press: New York, NY.

Data and other sources:

Borgatti, S.P., Everett, M.G. and Freeman, L.C. (2002). UCInet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies

FIFA (2003). *Regulations on the Status and Transfer of Players.* Available at: http://www.fifa.com/mm/document/affederation/administration/01/06/30/78/statusinhalt_en_122007.pdf (Accessed September 4th 2010)

Jousma, H. and Van Rossum, P. (2009). Science Based Business, Leiden University, Management Team Mobility data, personal communication.

REGPAT database, OECD, January 2010.

World Intellectual Property Organization (WIPO), International Patent Classification Guide, 2009. Available at: http://www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide_ipc_2009.pdf (Accessed September 26th 2010)