Microphysical proximity and knowledge flow on Bioscience Park Leiden

Pieter-Bart Visscher

MSc Thesis

Summer 2011

Microphysical proximity and knowledge flow

on Bioscience Park Leiden

---

name: Frederik Pieter Bart Visscher
born: 28-02-1986 in Utrecht
address: Korte Achterweg 7, 4458BG, ‘s-Heer Arendskerke
tel.no: 06 83 55 53 76
mail: pbvisscher@gmail.com

Master of Science Thesis

Erasmus University Rotterdam
Erasmus School of Economics
MSc Entrepreneurship, Strategy and Organisation Economics
First mentor: Dr. Sandra Phlippen

Delft University of Technology
Faculty of Architecture
MSc Real Estate & Housing
First mentor: Eng. Arch Herman Vande Putte MSc
Second mentor: Dr. Clarine van Oel

---

i Acknowledgements

This thesis could not have been realized without the patience, trust and support of the following persons, whom I am grateful for their contribution.

Firstly I would like to thank Sandra Phlippen for her straight-forward way of commenting on my writings, her ability to ask the necessary critical questions and her way of structuring my sometimes chaotic thinking, and of course for the coffee in the Urban Espresso Bar. Secondly I owe thanks to Clarine van Oel for her theoretical support in the statistical analysis. Although she was not entirely in favor of the method and application of Binary Logistic Regression she kept on stimulating to go in depth. Thirdly I am grateful to Herman Vande Putte for doing the exact opposite, namely to put my statistical findings into a broader perspective. He always brought my quantitative research into a qualitative context and showed in a thoughtful way his opinion on microphysical proximity.

I would also like to mention Harmen Jousma of Science Based Business Leiden, who gave some valuable insights in the daily practice of the Leiden Bioscience Park. Also Martijn Burger of the Applied Economics department of Erasmus University is worth mentioning, since he introduced the method Binary Logistic Regression.

Besides these academic supporters I would like to thank my parents, Frank and Ina Visscher, who for nearly seven years in a row not only contributed materially, but also mentally stood beside me. Without their help, trust and thrust I would not have come to this point. Finally I owe a great deal of thanks to Janine Visscher, my wife, from whom I asked a lot of sacrifices. Although not fully convinced of the added value of a double master degree she took care of all necessary tasks to keep me focused on studying, most recently in the care for our son Roemer.

Thanks to all of you!

ii Preface

This graduation project is a combined project for the Technical University of Delft and the Erasmus University Rotterdam. The chosen subject is of interest for both faculties, since it combines insights of both clustering theory as well as urban area theory. The chosen scale level is interfirm relations, which is not that common when looking at clustering effects. First and second mentor from the Technical University of Delft are respectively Herman Vande Putte and Clarine van Oel. Sandra Phlippen is the first mentor from Erasmus University. In an early stage of the project Harmen Jousma of Science Based Business of University of Leiden was involved.

This graduation project is part of a research project of the Erasmus University, in which several master students participate. All participating students followed the seminar Governance Clusters and Networks at the Erasmus School of Economics taught by Sandra Phlippen. The research project is loosely based on the clustering theories of the seminar. Prime focus of the research project is however one of the main case studies of the seminar; the Leiden Bioscience Park. Although students chose related subjects, main advantage is joint data gathering.

This thesis subject is of interest since it sheds light on the relation between the tangible factor physical proximity and rather intangible factor knowledge flow. Several theorists have stated the positive relation between proximity and knowledge flow, however the quantitative substantiation of this relation is limited. Since there tends to be a strong relation between knowledge flow and innovation, this deficiency of empirical foundation is peculiar. Fully understanding the factors that induce knowledge flow and thus innovation could be very beneficial since long term economic growth is largely dependent on innovation.

Contents

1. Introduction - Microphysical proximity and knowledge flows on Bioscience Park Leiden

The main focus area of this study is the Leiden Bioscience Park. This science park is a designated area near the central station of Leiden. It comprises more than sixty organizations involved with bioscience, among which the University of Leiden, the Leids Universitair Medisch Centrum, and several privately and publicly held companies. It is often stated that by bringing these companies together innovation is stimulated. The official website of the Leiden Bioscience Park puts it like this:

> "In today's global and knowledge-intensive industries, a company **cannot survive alone**. **Proximity to**, and **interaction with** other companies and knowledge institutions have become as important as financing, facilities and a good business climate." (bold added, source: official website LBP 2011)

In this study the effect of the factor 'proximity' on the flow of knowledge is studied. Noting that the flow and creation of knowledge is of vital importance to companies within knowledge intensive industries, companies should be searching for the best possible locational context. This study discusses the relation between their microphysical environment and its relation with knowledge flows. In this study proximity is treated as the inverted distance between two companies. Knowledge flow is defined as the flow of *fluid* knowledge between two companies in the form of co-patent production, labour mobility and/or joint research programmes.

The main findings could lead to an increased awareness of the geographical component of clustering theory. In recent years a vast amount of attention went to the relation between information and communication technology in relation to knowledge flow. The field of physical proximity and knowledge flow laid relatively dormant. This thesis could shed new light on the issue by using an elaborate database to analyze the relationship.

This study can be of interest for several parties. Urban planners and policy makers might want to use the findings as input for future urban planning processes. The findings might thus function as the evidence in evidence based design. Furthermore the results could be of interest to the academic world since this study focuses on the scale level of businesses, whereas clustering theory primarily is focused on the scale level of areas.

2. Leiden Bioscience Park: Introducing the area

The Leiden Bioscience Park is a designated area to the west of the city centre of Leiden. The area is especially dedicated to firms in the bioscience sector by the local and regional governments. This means that companies that are not involved with biotechnology or biopharmaceutical production are not allowed to settle on the park. Around sixty companies are located in the park, the majority of which have biopharmaceutical activities such as drugs and diagnostics research. The park covers an area of approximately 75 hectare - roughly 400 times 1800 meters. Figure 2.1 shows the location of Leiden within the Netherlands. Leiden is situated in the Western part of the Netherlands in the polycentric area named Randstad. Leiden is relatively close to Rotterdam, Amsterdam and The Hague. Figure 2.2 shows the situation within the municipal borders of Leiden. The Bioscience Park (yellow) is located close to the city centre (red) and is adjacent to the central station.

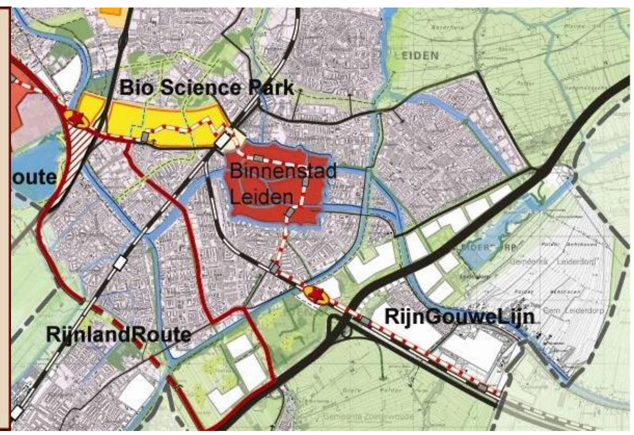

Figure 2.1: Leiden in the Netherlands. (Retreived from University of Maryland http://www.international.umd.edu/Infrastructure)

Figure 2.2: Bioscience Park in Leiden (Retrieved from http://ro.leiden.nl/planproces/Planproces/plannen/)

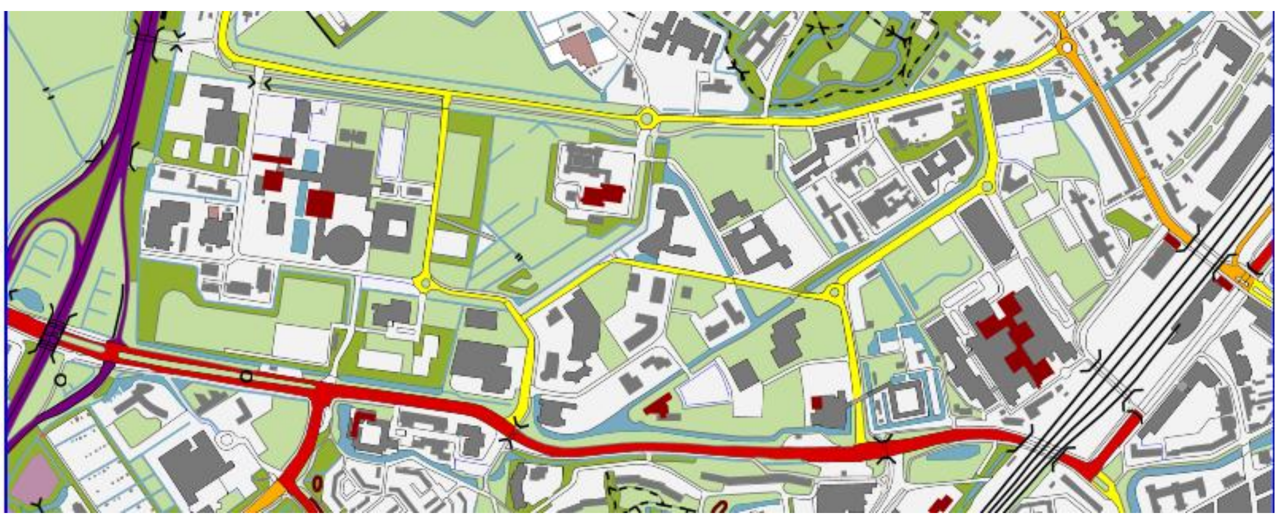

Figure 2.3: Leiden Bioscience Park (map via Geoloket on kaartenkamer.library.tudelft.nl)

Figure 2.3 shows the bioscience park and its geographically borders. These borders are:

- North: Wassenaarseweg (in yellow)
- East: railroad connection Den Haag – Schiphol (in black)
- South: Plesmanlaan (in red)
- West: highway A44 (in purple)

The bioscience industry is characterized for its knowledge intensity. Developing an innovative drug comes with high costs, years of investment to come to a patentable product from which benefits can be extracted for many years. The knowledge level required for the innovation and production is high as it involves clinical testing, specialized quality checks, manufacturing and marketing of the product.
The Leiden Bioscience Park plays an important role in the world of bioscience. It is generally known by the public for the world's first genetically modified cow ("Herman"). The most important discoveries made by business on the park are:

* Medicin against rheumatic arthritis (2mln produced) by Remicade
* World first fluid vaccine Quivaxem (500mln produced) by Crucell
* First medicin against Duchenne muscular dystrophy (fase 3) by Prosensa and the LUMC
* Medicin against Cockayne Syndrom / Progreria by DNAge

Figure 2.4 Aerial shot of Leiden Bioscience Park (retrieved from www.indodutchconnect.com/articles/leiden-bioscience-park_84.html)

Figure 2.4 shows the Bioscience Park area shown from the east. In the front of the picture the railroad track and the Leiden Central station are shown. Right behind the Central Station the Leiden University Medical Centre is situated (beige/yellow cubic shaped building). Behind the LUMC the companies are situated. According to calculations of the Bioscience Park foundation there are around twelve thousand people working on the Leiden Bioscience Park. Figure 2.5 and 2.6 give an overview of the division of these people over the different sectors and companies. As can be seen, the educational institutions and the LUMC employ the majority of the people in the lifescience industry. What should be noted however is that these figures (in table 2.5) are somewhat inflated for marketing mechanisms. In this study a stricter definition of 'working on the park' is used. This still results in a total amount of nearly ten thousand employees.

Company name	No. of employees
Life science companies	**Ca 3000**
Leiden University Medical Centre	**Ca. 7000**
Educational Institutions	**Ca. 2000**
Total	**Ca. 12000**

Figure 2.5: employees on LBP (Source: Factsheet LBP 2010 retrieved from www.leidenbiosciencepark.nl

Company name	Size	Company name	Size	Company name	Size
Add2xBio	7	EJR Quartz	6	Promasys	3
Aeon Astron Europe B.V.	2	Enzyscreen	2	Prosensa	35
Amarna Therapeutics	6	Flexgen	4	Protein Labelling Innovation	2
Apotex - Leiden	180	Fytagoras	1	Proteonic	2
BAC B.V. R&D	25	Galapagos	31	Proxy Laboratories	32
BaseClear	20	Genencor BV	74	Questions, Answers and more	6
Batavia Bioservices BV	12	Gimaris	2	Rahu Catalytics	7
Biocult	3	Giotto Management Consultants	2	Rainbow Oxidations	1
Bioke	30	Hal Allergy	40	ROC-Leiden Laboratoriumtechniek	20
Biotop medical	5	ISA Pharmaceuticals	3	Service XS	7
Cam Bioceramics	49	LAP&P	12	Servier Nederland Farma BV	88
Centocor	989	Leiden Probe Microscopy	3	STI Management	1
Centre Human Drug Research	99	Leiden University	1008	Ti-Pharma	11
Cosine	20	Leidse Instrumentmakersschool	29	TNO - Kwaliteit van Leven	77
Crucell	325	Isj medisch projectburo	3	To-BBB	6
Culgi	13	LUMC	6526	Verilabs	3
CWTS bv	16	Mentor	181	Viruvation	3
Deltacell	5	MEVS	3	Xendo	65
Derphatox	1	Octoplus	63	ZF-Screens	4
DNage	5	Pharming	44	ZoBio	6
Dutch Space	271	Profibrix	5		

Figure 2.6 Table of companies present on the Leiden Bioscience Park

3. Theoretical Framework

Separating working from living

The tendency to separate working from living environments is mainly a phenomenon from the last century. Combining the two functions of working and living within one location was rather common for craftsmanship and even for smaller agricultural activities. However since the Industrial Revolution there is a tendency to separate living and working area's more strictly. The clear separation was driven by several factors. Not only the increased demand for space played a role (due to scaling possibilities), also the harmful effects of production (negative externalities) had its influence. However one of the main drivers is to be found in the divergence of wishes for the surroundings of the buildings:

> "Whereas residential developers might tout their proximity to parks and churches, the field of industrial dreams offered its own set of amenities in infrastructure, services, and distance from city taxes and nuisance ordinances" (Vitiello 2005 p256)

This citation makes clear that the surroundings are of considerable influence for the separation of work and living when it comes to the industrial side of working.
However, already in the 16th century the first known example of an office was introduced in Florence. Architect Vesari was ordered by Cosimo de'Medici, duke of Florence, to draw and construct a building that housed the offices of the Magistrates and Guilds that governed the city. The building was called Uffizi, which was the base for our word office (MIBAC 2008).

Benefits of business parks

Already at the end of the 19th century there were theoretical confirmations of the benefits of separating work and living areas. Not the separation had a positive effect however, merely the grouping of similar activities had beneficial effects. In his seminal writings on internal and external economies, Alfred Marshall states that there are benefits to geographical co-location for businesses in the form of labour pooling of specialized workers, specialized inputs from suppliers and business related knowledge flow (1890 and 1920). In the end of the twentieth century and the beginning of the twenty-first century a vast amount of literature on business parks sums up the benefits of co-location (see for instance the work of Michael E. Porter). The benefits mainly boil down to economies of scale, reduced search costs, infrastructural facilities, labour pooling and shared services. A TU Delft graduation study with respect to the Leiden Bioscience Park also shows similar benefits (Van den Bergh, 2005). When asked what the major benefits are of being situated on the park, businesses mainly point at reduced costs due to diversification. A similar observation makes Nettie Buitelaar, Managing Director of the Leiden Bioscience Park Foundation when she states that companies benefit from each others knowledge of 'market introduction, regulatory affairs, logistics, quality control and assurance, clinical trials, etc.' (CONNECT 2010)

Business parks and science parks

The characteristics and beneficial conditions not only count for business parks, they also count for science parks. The fundamental difference between business parks and science parks is the nature of their business. Where business parks mainly consist of businesses engaged in commercial activities, science parks are aimed at business in knowledge intensive industries. When looking at these knowledge intensive industries, such as the bio and life science sector, another production factor besides land, labour and capital comes at play. This production factor is knowledge. Knowledge is a driver for innovation.

Conditions for innovation

Access to knowledge and knowledge flow are of crucial importance to companies in this field since they offer opportunities for innovation. Breschi and Malerba – when discussing innovation in knowledge intensive industries - state it as follows; "(…) performance ultimately depends on a set of factors and resources – knowledge, capabilities, skilled human capital, institutional and organizational structures (…)" (2005 p1). Access to knowledge is thus one of the key ingredients for successful innovation and consequently firm performance. However, according to Freeman (1987), and many other theorists that build upon the concept of innovation systems, innovation is a process done collectively. The process of innovation requires interaction between several firms and organizations, such as research centers, governmental bodies, universities. Individual firms in this collective learn by "embracing user-producer relationships, formal and informal collaborations, interfirm mobility of skilled workers and the spin-off of new firms from existing firms, universities, and public research centers". (Breschi and Malerba 2005 p3)

Forms of knowledge

The practice and theory of innovation shows us that firms engaged with innovation use a typical form of knowledge. They are working with uncodified knowledge that is hard to transfer. Knowledge stays liquid when the innovation process is performed. Polanyi was the first to give a name to this liquid knowledge and his expression was used ever since. He described it as 'we can know more than we can tell' (Polanyi 1966). The name given to this sort of information was tacit knowledge. It differs from codified knowledge, which is formalized knowledge that can easily be transferred in a depersonalized way by for instance patents, publications, operating manuals.
There are several factors that play a role in the distribution of tacit knowledge. These are physical encounters, absorptive capacity and low-tech learning.

Physical encounters

"Being personal or context-dependent, tacit knowledge represents disembodied knowhow that is acquired directly through interactive learning" (Howells 1996 p95). As Howells states, this tacit knowledge flows through interactive learning in which of course the frequency of interaction plays an important role. This process of interactive learning is sometimes referred to as learning by doing.

Absorptive capacity

However not only repeated physical encounters are a factor in the flow of tacit knowledge. The factor of cognitive proximity plays a similar important role. With cognitive proximity is meant the ability to understand what is transmitted. So not only the access to knowledge (through face-to-face contact) also the ability to comprehend what is transmitted. This notion is also reflected in the following quote, where it is called absorptive capacity:

> "Since knowledge creation and learning often depend on combining diverse, complementary capabilities of heterogeneous agents within and between organizations (Nooteboom 2000), there is strong need to bring these together. This is, however, not easy to do. The tacit and idiosyncratic nature of much knowledge implies that access to relevant knowledge is not a sufficient condition. The effective transfer of knowledge requires an absorptive capacity to identify, interpret and exploit the new knowledge (Cohen and Levinthal 1990)" (Boschma 2005 p63)

Low-tech learning

There is however a type of knowledge creation that is as influential as the earlier mentioned 'high-tech knowledge', however it does not take the 'dedicated investments' into university educations and research programs. This low-tech learning as Laestadius (1996) and also Maskell (1998) describe it, is an important driver of learning and innovation. Although these blue-collar workers are not involved in the direct production of the knowledge, their ability to handle, apply and use the knowledge is of invaluable effect to the conditions of knowledge flow. Due to their everyday handling of the knowledge and the structuring of the business practices, logistical services, resource management, sales, industrial relations (see for instance Malerba 1992) the creation of knowledge is supported.

Physical proximity and the flow of tacit knowledge

As we have seen innovation and firm performance are positively linked to knowledge and knowledge flow. Several studies have shown in this respect that the flow of knowledge (and more specific the externality of knowledge spillover) is in a way related to physical proximity. Breschi and Malerba express this quite clearly when they say that new knowledge is more easily and more efficiently spread between actors that are closely located (2005). Although the concept of localized knowledge flow received little attention in the years following Marshalls initial statements, since the 1990's renewed attention is given to the relation between geographical proximity and knowledge flow. Audretsch and Feldman (1996) discuss the findings of Marshall and Krugman in the following way:

> As Alfred Marshall (1920) and, later Krugman (1991) argue, there may be geographic boundaries to information flows or knowledge spillovers, particularly tacit knowledge, among the firms in an industry. Although the cost of transmitting information may be invariant to distance, presumably the cost of transmitting knowledge rises with distance. That is, proximity and location matter. (Audretsch and Feldman 1996 p630)

Also Hervas-Oliver and Albors-Garrigos (2009) discuss this relation and summarize that localized knowledge flow entails that innovation is closely connected to distance and location in geographically clustering of economic activities. In line with these findings is the study of Morgan (2004) when he evaluates the role of geographical proximity and more specifically tacit knowledge. He indicates that tacit knowledge is geographically bound since this knowledge type is highly dependent on the context and tied to persons, explaining why knowledge intensive activities tend to geographically cluster.

Tacit knowledge and ICT
Morgan continues in saying that although more and more knowledge intensive activities are taken place digitally, the ongoing digitalization cannot be considered a viable substitute for geographic proximity. In his study he explains that physical encounters are of utmost importance for the flow of tacit knowledge. He states that (virtual) proximity made possible through modern communication systems will never be a genuine surrogate for geographical propinquity. As he explains the complex nuances of non-verbal communication can only be transmitted in face-to-face contact, and can hardly be substituted by electronic means.

Direct role of physical proximity towards tacit knowledge
Sorenson (2005) comes with a supplementary explanation for the fact that firms within the same industry tend to co-localize. He argues that geographical proximity facilitates the frequency of interaction in a social network, and he indicates that this network is physically limited. In the article he first states the rational argumentation for the decision where to locate. As Sorenson makes clear one might expect that choosing a location depends highly on business economical motivations such as reducing the costs of transportation. A rational evaluation of all different possible locations and a logical process for selecting the best geographical place is what one anticipates. However the choice of a location is far less sophisticated. He continues in saying that reasons for co-location are far more down to earth and seldom involve rational decision making, Sorenson states that founders of new firms simply exhibit geographic inertia – they nearly always start their firms in the same communities in which they have been living and working (Sorenson 2005 p298) Personal reasons thus overshadow rational choices.

Definition of proximity
What does proximity mean? Jonsson (2002, see Moodyson & Jonsson 2007)) provided an overview of nearly all proximity definitions in the literature on clusters, districts and innovation systems and showed the ‘considerable elasticity in the notion of proximity’ (ibid. p117). He indicated that ‘the concept is used to encompass anything from a science park, or a city region to a whole continent’ (ibid. p117). Jonsson herewith showed the wide variety in the notion of proximity. In nearly every case ‘proximate’ can only be defined in relation to somewhat else. For example: when looking at the national level, companies situated in the Randstad conurbation are proximate.

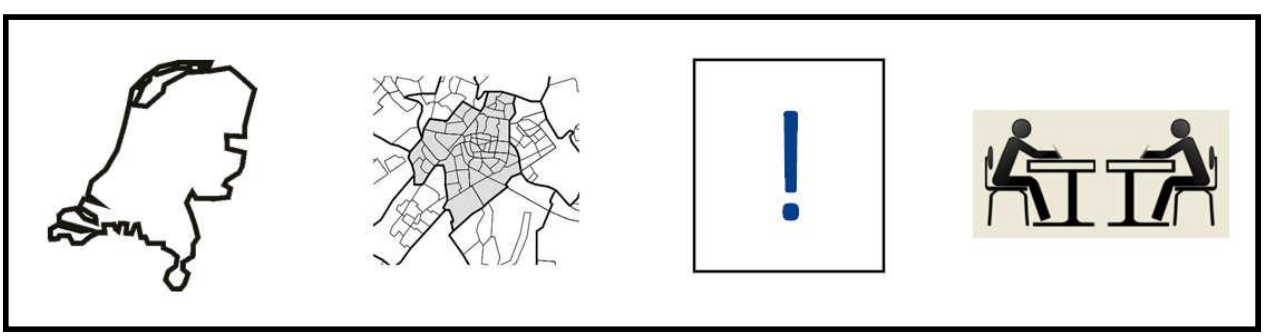

Figure 3.1: the scale levels of proximity (own production – exclamation mark indicates the chosen scale level)

- When looking at a national level, companies situated in the Randstad are proximate.
- When looking at the Randstad level, companies situated in the Leiden region are proximate.
- When looking at the Leiden region, companies situated in the Leiden Bioscience Park are proximate.
- When looking at the Leiden Bioscience Park, do more proximate companies have more knowledge flow than less proximate companies?

Proximity in this study

Being proximate to each other in this study is being situated on the Leiden Bioscience Park. This first layer is obvious Within this scale level a further distinction is made between companies that are more proximate to each other. To explain this graphically, company [A] is more proximate to company [B] than to company [C]. All three companies [A,B,C] are however proximate since they lay within the boundaries [bold black borders] of the designated area.

Separation between physical and microphysical proximity

This separation between scale levels gives insight in how proximity actually works and how it is related to knowledge flow. Exactly knowing whether the scale level of microphysical proximity is of importance would give rise to increased attention to the spatial layout of science parks. Since there is a clear (theoretical) link between economic growth, innovation and knowledge flow, it is of importance to know which factors induce the flow of knowledge. If microphysical proximity plays a large role, the locational choices for businesses would be made in a whole different manner. Rental prices for centrally situated buildings would go up, since the chance of being successful would increase. Also investors would be more likely to invest in the epicentre of knowledge flow. All these insights lead to the formulation of a main question:

Main question:

> **Is there a significant and positive relationship between microphysical proximity and knowledge flow?**

In essence this means the same as what is depicted in figure 3.2: is company [A] more likely to engage in knowledge flow with company [B] than with company [C]?

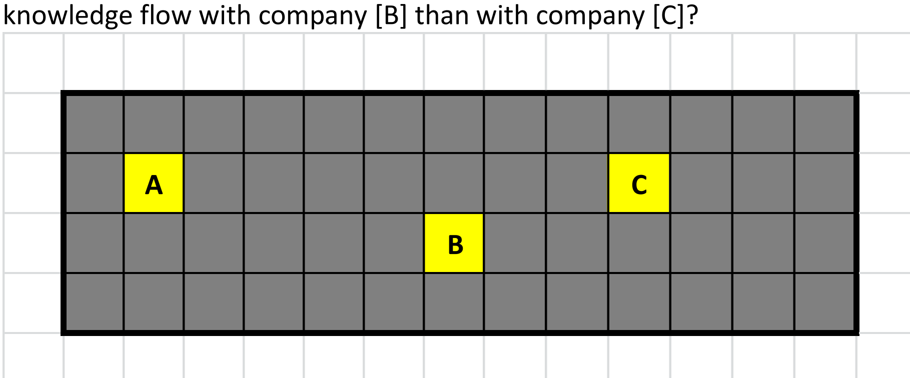

Figure 3.2 Is company A more likely to engage in knowledge flow with B than with C?

4. Operationalization

Proximity

Proximity is generally defined as the state of being near to something else. As indicated proximity is a relative concept. Notwithstanding there is a rather straightforward way in which proximity can be expressed, namely distance (expressed in meters). In this study geographical distance is taken as a measure of proximity. Distance is a numerical expression of spatial separation. In this study the distance between firms is measured between the centre of gravity of the buildings they are accommodated in.
Microphysical proximity is thus treated in this study as (inverted) distance between companies. When distance increases proximity decreases.
In order to measure the distance between two companies, the Euclidian distance between the centres of gravity are computed by using their exact GPS coordinate. These coordinates were retrieved by using their address and the GPS function in Google Earth. This results in a list of 62 longitudinal and latitudinal coordinates. These coordinates were converted to radians, and Excel computes the distance by using the command:

```
=ACOS(COS(RADIANS(90-A2)) *COS(RADIANS(90-A3)) +SIN(RADIANS(90-A2))
*SIN(RADIANS(90-A3)) *COS(RADIANS(B2-B3))) *6371
```
[1]

When comparing the outcomes of this distance computation with manual measurements (ruler mode) of Google Earth it turns out that the maximum difference between computed and measured is around 5%.

Knowledge flow

Where the concept of microphysical proximity can be seen relatively easy as inverted distance, the concept of knowledge flow is somewhat more difficult to define. In an earlier mentioned quote of Breschi and Malerba it was stated that individual firms *learn* through several different means. They mentioned user-producer relationships, formal and informal collaborations, interfirm mobility of skilled workers and the spin-off of new firms from existing firms, universities, and public research centers (Breschi and Malerba 2005 p3). Burger et al (2009) talk about knowledge spillovers and come to the following conclusion; “one can distinguish between at least three forms of knowledge spillovers (Boschma & Frenken 2006): spinoff firms, labour mobility and R&D collaboration.” (Burger et al. 2009 p141). In addition to this summation co-patent production should be mentioned, since research on the diffusion of knowledge in geographical areas has shown that also co-patent production is a form of knowledge flow. Jaffe et al (1993) describe this as follows:

---

[1] Retrieved from http://www.movable-type.co.uk/scripts/latlong.html and http://bluemm.blogspot.com/2007/01/excel-formula-to-calculate-distance.html

> “But knowledge flows do sometimes leave a paper trail, in the form of citations in patents. Because patent contain detailed geographic information about their inventors, we can examine where these trails actually lead.” (Jaffe et al. 1993 p578)

Thus co-patent production can also be seen as a way in which knowledge flows between companies. These findings lead us to the following list of *official* knowledge flow.

1. Co-patent production
2. Labour Mobility
3. Official Partnerships
4. Spin-offs & Start ups

Note that in this list there the fourth form of knowledge flow is a combination of both spin-offs and start ups. The main difference between the two is that in the case of spin-offs a mother company is involved, whereas in a start up there are no official ties with another company. However, as Sorenson showed in his study (2005) there are unofficial ties to other companies, since entrepreneurs that start up a business are geographically bound due to their social network.

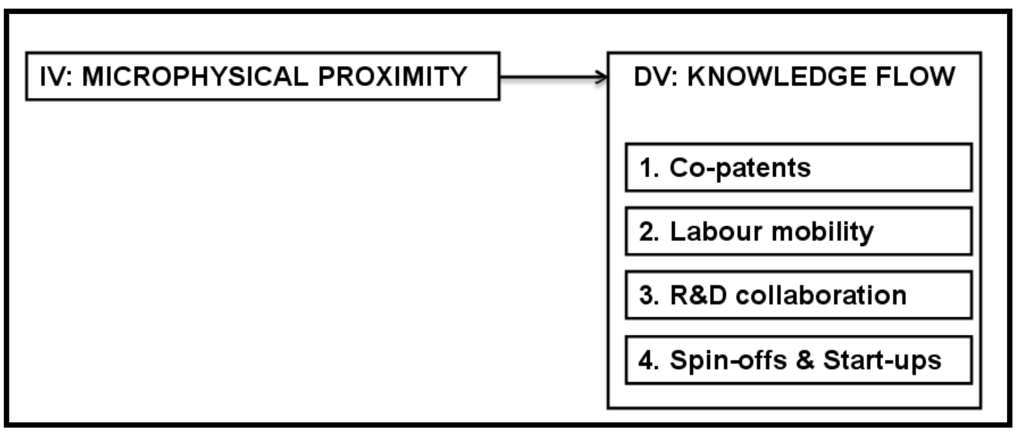

Figure 4.1: Graphical overview of the study (own production)

1. Co-patent production

Producing a patent in cooperation with another firm is a clear indicator of successful connection between firms. Linking connection co-patent production to micro geographical proximity is rather easily done by using a technique that is used in explorations of knowledge dispersion. This technique is used to assess knowledge flows by using patent citations.

2. Labour mobility

Labour pooling is seen as one of the main advantages of cluster formation. There is thus reason to look at the mobility of employees of bioscience companies in the Leiden bioscience area. A database from the Science Based Business Centre of the Leiden University is available for this analysis. This database contains the job history of 163 members of Leiden bioscience management team members. Knowing where the companies were based during the hopping of the managers thus gives information on the geographical distance between the company they left and entered.

3. Official partnerships

Official partnerships is a clear indicator of companies working together. These official partnerships, for instance in joint research and development, are published in annual reports and usually also on companies websites. There are however several limitations in the data gathering, since not every company is obliged to publish an annual report (depending on size and juridical structure), and websites do usually only contain current partnerships. In the data gathering period it shall become clear to what extent the data is available and to what extent is it useful in this study.

4. Spin-offs and start ups

There are two relations that have particular attention; firstly the distance between the newly erected company and its alma mater and secondly, when more than one founding partner is involved, the distance between these founding partners. Another database of the Centre for Science Based Business contains all information that is needed to conduct this analysis. This database contains a list of all companies that entered and left the park since the establishment of the park in 1985. The type of entry (being startup, spinoff, move from elsewhere etc.) is known and for the majority also its location within the park.

As stated the aim is to study whether microphysical proximity is of significant influence on knowledge flow within the Leiden Bioscience Park. For the statistical testing of the relationship a regression analysis is used. The independent variable within this study is microphysical proximity and the dependent variable is knowledge flow. However as indicated there are more factors that could have an effect on the chance of knowledge flow, these are size difference, age difference, and value chain difference. These factors will be added as covariates to the study. The basic structure of analysis is graphically:

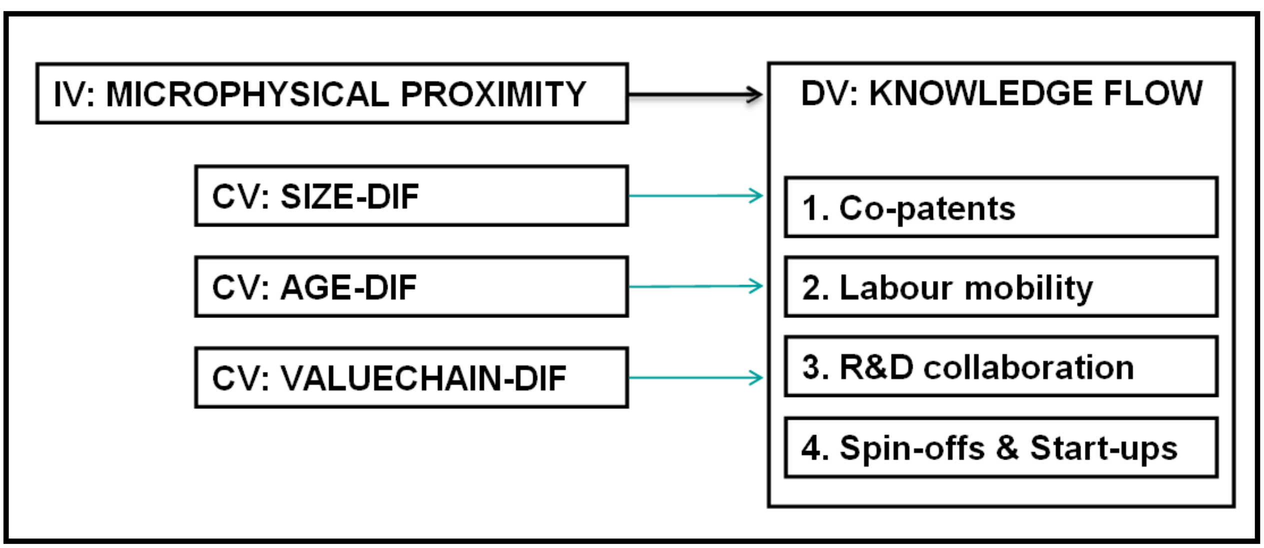

Figure 4.2: Graphical overview of the study (own production)

In this analysis we are comparing the companies that have experienced knowledge flow with companies that did not experience knowledge flow. What the analysis should tell us is whether microphysical proximity explains part of the observed knowledge flow links. It should show if companies that are more proximate to each other (with respect to other companies also on the park) have a higher probability of forming a knowledge flow link. Since there are 62 companies of interest on the LBP there are 1891 possible knowledge flow links within the science park:

$$(62*61)/2 = 1891$$

Since we are dealing with a limited dependent variable not every type of regression analysis can be used. Limited means that the dependent variable either is non-numerical in nature (categorical, ordinal) or that it can only take a restricted amount of discrete values. In this study the dependent variable can only take two values, being: knowledge flow (1=yes, 0=no). This requires a special type of regression analysis, binary logistic regression is used most often in this case.
The data needed to perform the statistical analyses as proposed in the operationalization is not fully available. There are however several databases concerning distinct aspects of the businesses on the park. The available data will be described first after which the combined dataset is introduced.

This database was converted into a dataset comprising 1891 cases. With the initial information the extra variables distance (DISTANCE), size difference (SIZEDIF), age difference (AGEDIF), and value chain difference (VCDIF) were computed. How this is done is explained hereafter:

- SIZEDIF

  The difference in size expressed in number of employees. Easily computed by:

  =ABS(SIZE company A – SIZE company B)

- AGEDIF

  Age difference between companies. Since we are looking at the relational dimensions of two companies, we do not merely adjust for company specific aspects. The age difference is computed by:

  = ABS(AGE company A – AGE company B)

- VALUECHAINDIF

  In order to get insight in the competitive forces between companies on the LBP we need to determine in which part of the value chain they are active. Fortunately the LBSPF has created an online Google Map connected database that classifies each company. A screenshot from the database is shown in figure 9. The companies can be placed in six categories, namely:

  1. Discovery and Research
  2. Clinical Trials
  3. Manufacturing

4. Distribution
5. Product & Process Development
6. Service, Sales and Other

This is of course congruent with the actual development process in the bioscience industry. A company which is placed in the categories 1,2 and 3 has a value chain difference of “2” with a company which is placed in categories 5 and 6. Note that the minimal distance is computed, not the maximal distance. (5-3 instead of 6-1). If the value chain difference is “0” the companies are competitors, if the difference is “>0” they are potential co-operators. The computing of the difference is done manually.

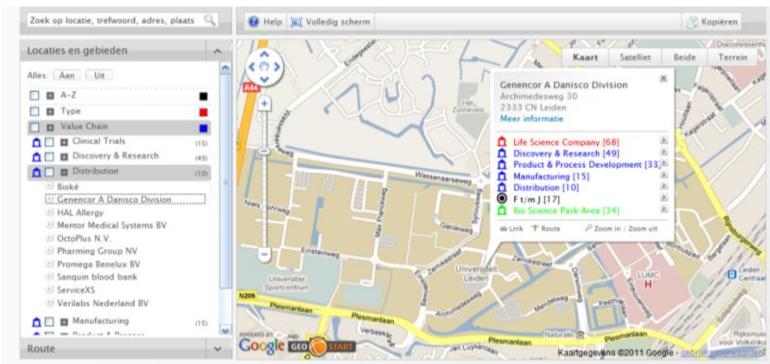

Figure 4.3: the online database of the LBSPF

The dummy variables startup/spinoff and incubator were also inserted. To this database four binary variables were added. These variables are:

- PATENT (whether companies created a patent together)
- LABMOB (whether an employee has switched from one to another companies)
- RDCOOP (whether the companies engaged in joint research contribution)
- KNOWFLOW (a variable that indicates whether there was any kind of knowledge flow)

The information from patents was retrieved via Espacenet which is the only search function of the European Patent Office (http://t1.espacenet.com/), the source of labour mobility was already explained. R&D cooperation was retrieved by working through annual reports (if present), official statements and 62 company websites.

Subquestions

This results in the following subquestions that are all addressing one of the success indicators:
1. What is the relation between microphysical proximity of firms and their co-patent production?
2. What is the relation between microphysical proximity of firms and their labour mobility?
3. What is the relation between microphysical proximity of firms and their R&D cooperation?
4. What is the relation between microphysical proximity of firms and their spin-offs and start ups?

Additional hypotheses
It is interesting to formulate some hypotheses to test for expected outcomes. Two hypotheses are constructed, the first regarding hierarchies in knowledge flow, the second regarding competition seen in knowledge flow.

Hypothesis 1
It would be unlikely that microphysical proximity has exact the same influence on every different type of knowledge flow. It is thus likely that we observe different relations between the proximity and the flow of knowledge. There could thus be some kind of hierarchy in forms of knowledge flow. If microphysical knowledge has a strong influence on the spread of tacit knowledge we could similarly argue that microphysical proximity thus has the strongest effect on the type of knowledge flow that involves the largest amount of tacit knowledge. This differentiation can be compared to the economic concept of price elasticity; a change in price results in a change in demand for a product, however the magnitude of this effect is not for every product the same. The suggested hierarchy would be:

1. Labour Mobility
2. Spin-offs & Start ups
3. Co-patent production
4. Official Partnerships

The H0: the suggested hierarchy is significantly present on Leiden Bioscience Park
The HA: the suggested hierarchy is not observed on Leiden Bioscience Park

Hypothesis 2
In the theoretical framework we have assumed that the geographical proximity leads to increased possibilities to cooperate. However, it might well be that this proximity does not merely lead to cooperation, but also to competition. A competitive environment would assume lower levels of cooperation and higher levels of competitive behaviour. Fierce competition might be reflected in a high level of labour mobility (companies buying strategic knowledge via employees) and startups and spinoffs (employees using knowledge from a company to start their own profitable venture) and lower levels of cooperation in patent production and official partnerships.
"Complementary signals scope for fruitful exchange while similarity in activities spells contest and market encounter. The firms in the vertical dimension of the cluster will, accordingly, often be business

partners and collaborators. The horizontal dimension will, on the contrary, consist mainly of rivals and competitors. (Maskell 2000 p 928)

It is thus of importance to include a variable in the analysis that deals with the position in the value chain to get an impression of these vertical and horizontal relations.

H0: Competitive forces are significantly present at the LBP
HA: Competitive forces are not significantly present at the LBP

The following two hypotheses deal with company's characteristics that might influence the observed relation between microphysical proximity and knowledge flow. Therefore the accompanying variables (size difference and age difference) are added in the regression as covariates.

Hypothesis 3

Regarding the size of the companies we could formulate the following belief. Although larger companies have more possibilities to engage in knowledge flow (e.g. more employees that could function as receptor), it is not unlikely that smaller companies are more likely to engage in knowledge flow. Since we are dealing with a knowledge intensive industry, small companies are more vulnerable and will more actively seek knowledge flow possibilities. Therefore the following hypothesis:

H0: smaller companies have significantly more knowledge flow than larger companies
HA: smaller companies do not have significantly more knowledge flow than larger companies

Hypothesis 4

Older firms have had more possibilities to engage in knowledge flow and might thus be more prominent in the study. As George and Zaheer (2004) put it: 'a firm tends to form more alliances over time and thus the older the firm the better its capabilities to generate knowledge flow." Therefore the following hypothesis:

H0: older companies have significantly more knowledge flow than younger companies
HA: older companies do not have significantly more knowledge flow than younger companies

5. Data

Binary Logistic Regression

This typical form of logistic regression should tell us whether the chance of a link (knowledge flow) is to be explained from the independent variables. It is popular since it overcomes some of the stringent assumptions of an Ordinary Least Square Regression. Firstly, Binary Logistic Regression (BLR) does not require a linear relation between the dependent and independent variable as OLS does. It can handle any kind of relationship since a log transformation (non-linear) to the predicted odds ratio is inserted. Secondly BLR can cope with nominal and ordinal data, and thus does not need metric independents (interval, ratio). Thirdly normality of the independent variables or the residuals is not necessary (although it might lead to more stable predictions). Fourthly the error terms (residuals) are assumed to be independent, which offsets the risk of autocorrelation. Fifthly the independents are not linearly related to one another which prevents the risk of collinearity. Sixthly homoscedasticity is not needed, BLR can handle heteroscedasticity for each level of the independent variable.

The BLR however does have some requirements. These are first of all that the dependent variable is binary, which is the case in this study. It secondly requires that the probability that knowledge flow (or any other event) occurs is set to “1” of the binary scale, since the model tries to predict P[Y=1]. This is covered in this study. Thirdly the model is rather sensitive to over- and underfitting and therefore as always a leanest model should be sought. This can best be done stepwise which will be done in this study. Fourthly the data should be independent, there should be no dependency in the error terms. Every observation should be unique and independent, which is the case in this study. Fifthly BLR requires the independents to be linearly related to the log odds, this is especially the case for metric data, such as the SIZEDIF, AGEDIF and DISTANCE variables in this study. Fortunately there is nothing troublesome to be seen from a discriminant analysis (see appendix). Lastly, and this is a drawback for the PATENT variable in this study is that BLR requires quite large sample sizes. In an OLS 5 observations per variable is the bare minimum, with BLR this lies somewhat higher at around 10. For the large majority of the data this is the case, however there are just 8 observations for the PATENT variable.

Since BLR requires one binary dependent variable, and we are testing multiple knowledge flow variables, Martijn J. Burger from the Applied Economics Department of the Erasmus School of Economics advised to insert the other dependent variables as independents in the model. So when testing co-patent production, the variables labour mobility, R&D collaboration and spinoff/startup are inserted in the model.

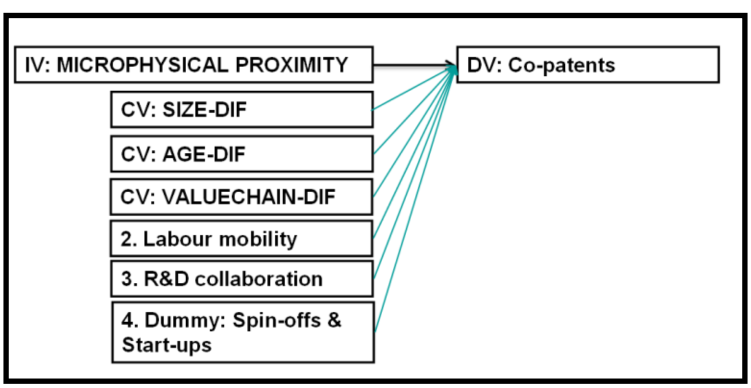

Figure 5.1: Graphical overview of statistical analysis

“Notwithstanding the intuitive awareness as to the rising importance of the theoretical debate on regions and clusters, there is a substantial lack of empirical support in ascertaining its precise magnitude and evolution.”(Cruz & Teixeira 2**0**09 p1264)

Existing datasets

The main database used in this study is a database provided by the Leiden Bio Science Park Foundation (LBSPF). This foundation is situated on the park and facilitates and supports startups as well as established companies on the LBP. The LBSPF started constructing a database for own use. This database contains information of 113 bioscience businesses that are situated in Leiden and surroundings. It thus also contains data of companies on the LBP. Although the database was far from complete, it included information on:

- year of entry on the park
- type of entry (startup, spinoff, relocation, etc.)
- year of exit
- type of exit (liquidation, relocation, takeover, etc.)
- number of employees

The missing values and replenished if possible with information from company’s websites, annual statements and the website company.info.

The second database was also initially constructed by the LBSPF and concerned the movement of top scientists and CEO’s on the park. The aim of constructing this database was to keep track of the labour mobility on the park. The data provided by the LBSPF contained 65 CEO and scientists. A student from the Erasmus University (Fleur de Groot) worked with this database for her thesis project and added information to it. In this study this database is used not to keep track of the employee themselves, however to get insight from which company to which company they shifted.

The basic database

From the existing datasets one dataset was created. Firstly all the company characteristics were compiled into one database. A number of 62 companies were left from the initial 113. The majority of the excluded companies was either not situated within the boundaries of the LBP or non-existent at this moment (taken over, merged, stopped to exist). From these 62 companies the following information was included:

- Company name
- Address (street, number, city)
- Latitude (in degrees)
- Longitude (in degrees)
- Size (no. of employees)
- Age (2010 minus year of establishment)
- Value Chain (position in valuechain, see further description below)

Dummy variables (0=no, 1=yes) on whether the company is a “startup/spinoff”, and whether the company was situated in an “incubator

6. Results

The results of the statistical analysis will be dealt with in separate subchapters. Firstly the Binary Logistic Regression will be discussed somewhat more elaborate.

6.1 Model overview

In this study four different models are presented, table 6.1 shows the models with the accompanying numbering. In Model 1 the aggregated knowledge flow variable is taken as dependent variable and four independent variables are added. The models 2, 3 and 4 have a particular form of knowledge flow as dependent variable; patent production, labour mobility and R&D cooperation respectively.

Each of the models has three submodels. The first submodels (e.g. 2.1, 3.1, and 4.1) consist of all effects, thus adding the variables sizedif, agedif, vcdif, and distance[2]. The second submodels (e.g. 2.2, 3.2, and 4.2) are the leanest models. In these the highest explanatory power with the least amount of variables is presented. In order to check the effect of distance (microphysical proximity) this variable is always added to the leanest model (stepwise, latest step). In the third submodels (e.g. 2.3, 3.3, and 4.3) the leanest model is supplemented with the other forms of knowledge flow. In the case of PATENT as dependent variable (model 2.3), LABMOB and RDCOOP are added as independent variable. In this way the effect of the different types of knowledge flow is studied.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Table 6.1 overview of all the statistical models

[2] The variable Spinoff/Startup produced no significant results in no single model, therefore it is left out completely out of the analysis.

6.2 Interpreting Binary Logistic Regression

In this part the outcomes of the Binary Regression Analysis are interpreted. For the explanation of the regression output several sources are used. The first source are the power points presentations provided by of Martijn Burger[3] has provided students in the seminar with several insightful PowerPoint presentations on Logistic Regression. The second source are the writings of Susan Collins. She has provided several Powerpoint as well as instruction videos on the topic of logistic regression and binary logistic regression[4]. Third source used as backup is the SPSS Results coach, which gives some interpretative guidelines, albeit quite limited.

The regression results are presented in one table, the full SPSS output per model is shown in the Appendix. Every table consists of three parts. Firstly the variables with their accompanying Beta, ODDS ratio and Significance are given. Secondly the output of the model fit test is given via -2LogLikelihood, Chi-square and p-value. Thirdly the explanatory power is shown via the Nagelkerke R-square.

The Beta and ODDS ratio are both an expression of effect size. The beta value is an ordinary regression coefficient and explains the size of the effect of the variable on the dependent variable. More telling however is the ODDS ratio, which is an Exponent(Beta). This statistic shows the likelihood of event-occuring and is centered around [1]. Values above 1 indicate a positive effect, values below 1 indicate a negative effect. An odds ratio of 1.400 thus means that

The Hosmer and Lemeshow Test is a Goodness-of-fit statistic that tests whether the null hypothesis is rejected or accepted. The null hypothesis assumes that there is no difference between the observed and predicted values (yes=1 and 0=no) of the dependent variable in the regression. If there indeed is no difference, the p-value should lie below the 0.05 threshold ($p<0.05$). However, in this regression analysis we are looking for the exact opposite, namely a significant difference between the (yes=1 and 0=no) values. Values of $p>0.05$ thus fail to reject the null-hypothesis and consequently show an acceptable level of model fit. The higher the p-value the better.

In an Ordinary Least Square Analysis the R-square statistic shows to what extent the variance in the dependent variable is causes by the independent variables in the model. The Cox & Snell R-square and the Nagelkerke R-square provide a similar insight, where the latter is also bound to values between 0 and 1. This Nagelkerke R-square is thus best comparable to the R-square statistic of OLS.

---

[3] Martijn J. Burger BA MSc is currently doing PhD research at the Applied Economics Department of the Erasmus University Rotterdam. His research interests are amongst others urban and regional economics, and spatial econometrics. He

[4] Susan E. Collins PhD is associate research professor of the University of Washington and should not be mistaken for Susan M. Collins – Republican Senator in Washington for the state of Maine.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
SIZEDIF	0.000	1.000	0.000
AGEDIF	0.066	1.068	0.000
VCDIF			0.034
VCDIF(1)	1.100	3.004	0.046
VCDIF(2)	-0.297	0.743	0.686
VCDIF(3)	0.310	1.363	0.677
VCDIF(4)	0.533	1.705	0.509
VCDIF(5)	-16.914	0.000	0.995
DISTANCE	-0.001	0.999	0.037
Constant	-4.375	0.001	0.000

-2 Log L	526.657
Chi-square	4.949 (df = 8)
P-value	0.763
NGK R2	0.138

This model has knowledge flow (KNOWFLOW) as dependent variable and SIZEDIF, AGEDIF, VCDIF and DISTANCE as independent variables. There are 65 cases in which there was knowledge flow (1=YES). Looking at the goodness of fit analysis it is clear that there is a more than adequate model fit. The explanatory power however is rather low, since the Nagelkerke R-square value is 0.138.

The variable SIZEDIF has a significant influence (p=0.000), however the effect is neutral (ODDS = 1.000), so nor positive nor negative. The variable AGEDIF also has a significant influence (p=0.000), the effect of this variable is positive (ODDS is 1.068). This means that a larger age difference (increases with 1 year) leads to a 1.068 larger chance of engaging in knowledge flow. The variable VCDIF also has a significant influence, however this is limited to VCDIF(1) (p=0.046). The effect however is considerable (ODDS = 3.004) meaning that companies with a value chain difference of one are three times more likely to engage in knowledge flow than companies without a value chain difference.Looking at the variable distance we see that it has a significant effect (p=0.037) however that the extent of the effect is very limited (ODDS = 0.999). The constant is negative (β = -4.375) and significant (p=0.000).

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
SIZEDIF	0.001	1.001	0.000
AGEDIF	-0.115	0.892	0.060
VCDIF			1.000
VCDIF(1)	16.858	2.1E+07	0.995
VCDIF(2)	0.708	2.030	1.000
VCDIF(3)	16.575	1.6E+07	0.995
VCDIF(4)	0.998	2.713	1.000
VCDIF(5)	0.884	2.421	1.000
DISTANCE	-0.001	0.999	0.543
Constant	-21.199	0.000	0.000

-2 Log L	87.965
Chi-square	5.587(df = 8)
P-value	0.693
NGK R2	0.235

This model has co-patent production (PATENT) as dependent variable and SIZEDIF, AGEDIF, VCDIF and DISTANCE as independent variables. There are 8 cases in which there was a patent produced by more than one firm (all situated on the park). Looking at the goodness of fit analysis it is clear that there is a more than adequate model fit. The explanatory power however is adequate, since the Nagelkerke R-square value is 0.235.

The variable SIZEDIF has a significant influence (p=0.000), however the effect is slightly positive (ODDS = 1.001). The variable AGEDIF does not have a significant effect, since the p-value (p=0.060) is just above the threshold. The effect would however be negative (ODDS = 0.892)if it were significant. The variable VCDIF does not have a significant effect (p > 0.05). Looking at the variable distance we see that it does not have a significant effect (p=0.543), the effect would be slightly negative. The constant is negative (β = -21.199) and significant (p=0.000). There are some consideration since the sample size (1=YES) is small (only 8), these will be given in the discussion.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
SIZEDIF	0.001	1.001	0.000
AGEDIF	-0.129	0.879	0.048
DISTANCE	-0.001	0.999	0.460
Constant	-4.787	0.008	0.000

-2 Log L	97.281
Chi-square	8.815 (df = 8)
P-value	0.358
NGK R2	0.152

This leanest model also has co-patent production (PATENT) as dependent variable and SIZEDIF, AGEDIF, and DISTANCE as independent variables. Again there are 8 cases in which there was a patent produced by more than one firm (all situated on the park). The goodness of fit analysis shows that there is moderate model fit. The explanatory power however is smaller than before, since the Nagelkerke R-square value is 0.152.

The variable SIZEDIF has a significant influence (p=0.000), however the effect is slightly positive (ODDS = 1.001). The variable AGEDIF has a significant effect, since the p-value (p=0.048) is just below the 5% threshold. Its effect is negative (ODDS 0.879 meaning that increased age difference decreases the likelihood. The variable distance does not have a significant effect (p=0.460). The constant is negative (β = -4.787) and significant (p=0.000).

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
SIZEDIF	0.001	1.001	0.000
AGEDIF	-0.145	0.865	0.035
DISTANCE	-0.001	0.999	0.456
LABMOB	0.653	1.921	0.631
RDCOOP	1.735	5.669	0.185
Constant	-4.787	0.008	0.000

-2 Log L	95.105
Chi-square	10.283 (df = 8)
P-value	0.246
NGK R2	0.172

In this model the leanest model is taken as basis and the other forms of knowledge flow are added as independent variable. So, PATENT is the dependent variable and SIZEDIF, AGEDIF, DISTANCE, LABMOB and RDCOOP are the independent variables. Again there are 8 cases in which there was a patent produced by more than one firm (all situated on the park). The goodness of fit analysis shows that there is moderate model fit. The explanatory power is not large, but increased somewhat, since the Nagelkerke R-square value is 0.172.
The variable SIZEDIF has a significant influence (p=0.000), however the effect is slightly positive (ODDS = 1.001). The variable AGEDIF has a significant effect (p=0.035) and has a negative effect (ODDS= 0.865) meaning that increased age difference decreases the likelihood of co-patent production by 0.865. The variable distance does not have a significant effect (p=0.460). Also the other forms of knowledge flow do not have a significant effect (p=0.631 and p=0.185 respectively). The constant is negative (β = -4.787) and significant (p=0.000).

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
SIZEDIF	0.000	1.000	0.004
AGEDIF	0.090	1.094	0.000
VCDIF			0.224
VCDIF(1)	1.742	5.709	0.098
VCDIF(2)	0.721	2.057	0.543
VCDIF(3)	0.835	2.306	0.509
VCDIF(4)	0.691	1.995	0.637
VCDIF(5)	-15.426	0.000	0.995
DISTANCE	-0.002	0.998	0.003
Constant	-5.619	0.004	0.000
-2 Log L	305.311		
Chi-square	7.168 (df = 8)		
P-value	0.519		
NGK R2	0.156		

This model has labour mobility (LABMOB) as dependent variable and SIZEDIF, AGEDIF, VCDIF and DISTANCE as independent variables. There are 33 cases reported of an CEO or CSO hopping within the bioscience park. Looking at the goodness of fit analysis shows us that there is an adequate model fit. The explanatory power however is low, since the Nagelkerke R-square value is 0.156.
The variable SIZEDIF has a significant influence (p=0.004), however the effect is exactly neutral (ODDS = 1.000). The variable AGEDIF also has a significant effect (p=0.000), the effect being positive (ODDS = 1.094). The variable VCDIF does not have a significant effect (p > 0.05 in all cases). Looking at the variable distance we see that it does have a significant effect (p=0.003) and slightly negative (ODDS = 0.998). The constant is negative (β = -5.619) and significant (p=0.000).

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
SIZEDIF	0.000	1.000	0.004
AGEDIF	0.099	1.104	0.000
DISTANCE	-0.001	0.999	0.002
Constant	-4.393	0.012	0.000

-2 Log L	319.958
Chi-square	6.735 (df = 8)
P-value	0.565
NGK R2	0.111

This leanest model also has labour mobility (LABMOB) as dependent variable and SIZEDIF, AGEDIF, and DISTANCE as independent variables. Again there are 33 cases in which there was a jobhop on the park. The goodness of fit analysis shows that there is adequate model fit. The explanatory power however is smaller than before, since the Nagelkerke R-square value is 0.111.

The variable SIZEDIF has a significant influence (p=0.004) however has a completely neutral effect (ODDS = 1.000). The variable AGEDIF has a positive (ODDS = 1.104) and significant effect (p=0.000). An increase in AGEDIF the likelihood of labourmobility 1.104 times. The variable distance has a significant and negative effect (p=0.002 and ODDS=0.999). The constant is negative (β = -4.393) and significant (p=0.000).

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
SIZEDIF	0.000	1.000	0.002
AGEDIF	0.079	1.082	0.022
DISTANCE	-0.002	0.998	0.001
RDCOOP	3.112	22.455	0.000
PATENT	1.276	3.583	0.299
Constant	-4.399	0.012	0.000
-2 Log L	282.385		
Chi-square	11.830 (df = 8)		
P-value	0.159		
NGK R2	0.224		

In this model the leanest model is taken as basis and the other forms of knowledge flow are added as independent variable. So, LABMOB is the dependent variable and SIZEDIF, AGEDIF, DISTANCE, PATENT and RDCOOP are the independent variables. Again 33 jobhops were observed. The goodness of fit analysis shows that there is low model fit. The explanatory power is not large, but increased somewhat, since the Nagelkerke R-square value is 0.224.

The variable SIZEDIF has a significant influence (p=0.002), however the effect neutral (ODDS = 1.000). The variable AGEDIF has a significant effect (p=0.022) and has a positive effect (ODDS= 1.082) meaning that increased age difference makes the the actors 1.082 times more likely to engage in labour mobility. The variable distance has a significant and negative effect (p=0.001 and ODDS=0.998). RDCOOP has a significant (p=0.000) and extremely positive effect (ODDS=22.455) This means that actors that engaged in R&D cooperation are 22 times more likely to also engage in labour mobility. Since this effect is large, a Crosstabulation with a Cramer's V and McNemar analysis is used to check whether these two variables are in fact different from each other. Cramer's V = 0.288 with a significance of (p=0.000) and McNemar Significance = 1.000 which indicates that there is not a troublesome form of autocorrelation. Only in 11 cases both RDCOOP and LABMOB are (1=YES).

Patent production does not have a significant effect (p=0.299). The constant is negative (β = -4.399) and significant (p=0.000).

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
SIZEDIF	0.000	1.000	0.161
AGEDIF	0.095	1.099	0.000
VCDIF			0.291
VCDIF(1)	0.542	1.720	0.393
VCDIF(2)	-0.831	0.435	0.373
VCDIF(3)	-1.044	0.352	0.377
VCDIF(4)	0.095	1.100	0.921
VCDIF(5)	-16.853	0.000	0.995
DISTANCE	0.000	1.000	0.733
Constant	-5.153	0.006	0.000

-2 Log L	323.496
Chi-square	6.673(df = 8)
P-value	0.572
NGK R2	0.122

This model has R&D cooperation (RDCOOP) as dependent variable and SIZEDIF, AGEDIF, VCDIF and DISTANCE as independent variables. There are 35 cases in which there R&D cooperation was observed. Looking at the goodness of fit analysis it is clear that there is an adequate model fit. The explanatory power however is low, since the Nagelkerke R-square value is 0.122.

The variable SIZEDIF does not have a significant influence (p=0.161). The variable AGEDIF does have a significant effect (p=0.000) The effect is positive (ODDS = 1.099) which means that a larger age differences increases the likelihood of R&D cooperation 1.099 times. The variable VCDIF does not have a significant effect (all p-levels are above the 5% threshold). Looking at the variable distance we see that it does not have a significant effect (p=0.733), The constant is negative (β = -5.153) and significant (p=0.000).

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
AGEDIF	0.111	1.117	0.000
DISTANCE	-0.002	1.000	0.540
Constant	-5.175	0.006	0.000

-2 Log L	341.345
Chi-square	5.775 (df = 8)
P-value	0.672
NGK R2	0.069

This leanest model also has R&D cooperation (RDCOOP) as dependent variable and AGEDIF and DISTANCE as independent variables. Again there are 35 cases in which there was a jobhop on the park. The goodness of fit analysis shows that there is adequate model fit. The explanatory power however is even smaller than before, since the Nagelkerke R-square value is 0.069.

The variable AGEDIF has a positive (ODDS = 1.117) and significant effect (p=0.000). An increase in AGEDIF increased the R&D cooperation 1.117 times. The variable distance does not have a significant influence (p=0.540). The constant is negative (β = -5.175) and significant (p=0.000).

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Variables

Variables	β	ODDS ratio	Sig
AGEDIF	0.095	1.099	0.000
DISTANCE	0.000	1.000	0.193
PATENT	1.898	6.671	0.103
LABMOB	3.118	22.607	0.000
Constant	-5.498	0.004	0.000
-2 Log L	301.174		
Chi-square	8.898 (df = 8)		
P-value	0.351		
NGK R2	0.188		

In this model the leanest model is taken as basis and the other forms of knowledge flow are added as independent variable. So, RDCOOP is the dependent variable and SIZEDIF, AGEDIF, DISTANCE, PATENT and LABMOB are the independent variables. Again 35 jobhops were observed. The goodness of fit analysis shows that there is adequate model fit. The explanatory power is moderate, and increased with respect to the leanest model, the Nagelkerke R-square value is 0.188.

1. The variable AGEDIF has a significant effect (p=0.000) and has a positive effect (ODDS= 1.099) meaning that increased age difference makes the actors 1.099 times more likely to engage in R&D cooperatoin. The variable distance does not have has a significant (p=0. 193). The variable PATENT does not have a significant effect (p=0.103). LAMMOB however has a significant (p=0.000) and extremely positive effect (ODDS=22.607) This means that actors that engaged in labour mobility are 22 times more likely to also engage in R&D cooperation. Since this effect is large, a Crosstabulation with a Cramer's V and McNemar analysis is used to check whether these two variables are in fact different from each other. Cramer's V = 0.288 with a significance of (p=0.000) and McNemar Significance = 1.000 which indicates that there is not a troublesome form of autocorrelation. Only in 11 cases both RDCOOP and LABMOB are (1=YES).The constant is negative (β = -5.498) and significant (p=0.000).

6.3 Overview results

The table in this section gives an overview of the results of the statistical analysis. In the first column the model number is shown. The second contains the dependent variable. The third column lists the independent variables. The colors of the independent variables show the effect on the dependent variable. Green represents a significant positive effect. Blue represents a significant negative effect, grey means either not significant or significant but with a negligible effect size. The outcomes are discussed in the next chapter

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, RDCOOP, PATENT
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

LEGENDA
SIGNIFICANT POSITIVE
SIGNIFICANT NEGATIVE
NON-SIGNIFICANT / SIGNIFICANT NEUTRAL

7. Conclusions

Looking at the results of the binary logistic regression several conclusions can be drawn. Firstly we will look at the data as they were presented in the results section and give an overview. Secondly the main question will be answered. Thirdly the hypotheses will be dealt with. Lastly an overall conclusion will be constructed.

7.1. Main Findings

The variable size difference is in all models of significant influence, however the effect is either positive or neutral. In general larger size differences increases the likelihood of engaging in knowledge flow. In model 4.1 the effect is temporarily gone (under the influence of other independent variables), however when these insignificant variables are removed the effect is clearly visible in the leanest models 2.2 and 4.2. The variable age difference has a positive and significant effect on the dependent variables labour mobility and R&D cooperation, and a negative effect on patent production. Looking at the overall model (model 1) the effect is positive. This indicates that with increasing age difference the likelihood of knowledge flow increases as well. This is however not the case for patent production.
The variable value chain difference only has a significant effect in the overall model (Model 1). In this model the effect is positive, meaning that knowledge tends to flow vertically, not horizontally. In both cases the influence is negative, which means that the smaller difference in activities of the value chain increases the likelihood of knowledge flow in the form of labour mobility. Employees thus tend to hop horizontally instead of vertically.
There is a strong link between the variables labour mobility and R&D cooperation. This effect is shown in the models 3 and 4. The effect is positive and significant in both directions. Meaning that R&D cooperation leads to labour mobility and that labour mobility leads to knowledge flow. More about this relation is stated in the discussion.

7.2. Answering the main question

In the introductory part of this thesis the main question is given. It sounds:

Is there a significant and positive relationship between microphysical proximity and knowledge flow?
The answer to this question is partly NO and partly YES. However, it should be noted that the question entails the expressions “significant” and “positive”. We could easily state that a significant relation between distance and knowledge flow is not found in the majority of the models. Only the model Labour Mobility shows a significant relation with distance. For all the other models the answer to the main question is NO. Looking at the Labour mobility model we can see that there is a negative relation between DISTANCE and LABMOB. This means that if distance increases, labour mobility decreases. Since microphysical proximity was defined as inverted distance (and thus distance can be seen as inverted microphysical proximity). There is evidence that if microphysical proximity increases, labour mobility increases. The answer to the main question for the labour mobility model is thus YES.

7.3. Answering the hypotheses

Besides the main question several hypotheses were formulated. These hypotheses were:

H1-0: the suggested hierarchy is significantly present on Leiden Bioscience Park
H1-A: the suggested hierarchy is not observed on Leiden Bioscience Park

H2-0: competitive forces are significantly present at the LBP
H2-A: competitive forces are not significantly present at the LBP

H3-0: smaller companies have significantly more knowledge flow than larger companies
H3-A: smaller companies do not have significantly more knowledge flow than larger companies

H4-0: older companies have significantly more knowledge flow than younger companies
H4-A: older companies do not have significantly more knowledge flow than younger companies

Tacit knowledge
H1-0: the suggested hierarchy is significantly present on Leiden Bioscience Park
H1-A: the suggested hierarchy is not observed on Leiden Bioscience Park

It was hypothesised that microphysical proximity had the largest influence on the variables that involved the largest amount of tacit knowledge. The suggested hierarchy therefore was:

1. Labour Mobility
2. Spin-offs & Start ups
3. Co-patent production
4. Official Partnerships

It is of course hard to compare the different models, however what could be said about the hierarchy is that microphysical proximity only plays a very minor role in the creation of knowledge flow. Therefore the differences between the forms of knowledge flow are also marginal.

Competitive forces
H2-0: competitive forces are significantly present at the LBP
H2-A: competitive forces are not significantly present at the LBP

There is only one clear indicator for the existence of competitive forces in this study. That indicator is value chain difference. In the overall model value chain difference plays a significant and positive role. This means that knowledge flow is more likely to occur in vertical (buyer-supplier) relations than in vertical (competitive relations). Since knowledge flow is a rather cooperative activity, this is telling for the competitive forces.

Company size
H3-0: smaller companies have significantly more knowledge flow than larger companies
H3-A: smaller companies do not have significantly more knowledge flow than larger companies

Size difference has in all models a neutral or positive and significant effect on knowledge flow. This should be interpreted correctly. The larger the size difference between two companies, the more likely they are in engaging knowledge exchange. However can this be explained by the eagerness of small companies or by the increased chance of larger companies? When looking at the table, it is shown that the average size of a company engaging in knowledge flow (231) is above the average size of all companies on the park (169) even when correcting for the outliers by using the median value, this is still seen. In all cases the median size of a company engaging in one of the forms of knowledge flow is larger than the median size of all companies on the business park.

SIZE	**AVERAGE**	**MEDIAN**
All companies	169	7
Knowledge flow	231	12
Patent	1132	201
LABMOB	296	16
RDCOOP	309	17

Table 7.1 comparing sizes

This would mean that the larger companies, among which the Leiden University and the LUMC, have a dominant position in the LBP, which is indeed the case when looking at the data.

Company age
H4-0: older companies have significantly more knowledge flow than younger companies
H4-A: older companies do not have significantly more knowledge flow than younger companies

The null-hypothesis is true. We have seen that age difference is positively related to knowledge flow in general and labour mobility and R&D cooperation. Only in the case of co-patent production the age difference is of significant negative influence.

8. Discussion

Interpretation

The results of this study show that the effects of microphysical proximity on knowledge flow tend to be negligible. The effect is either not significant or (when the effect is significant) it is extremely small. The obvious conclusion would be that this microphysical scale level is of no interest and could be left aside from now on. However this conclusion should be nuanced.

In this study the relation between knowledge flow and microphysical proximity is main point of interest. As seen, it is assumed that proximity (being situated on the Leiden Bioscience Park) is a beneficial condition. The question that arose was whether microphysical proximity played a significant role in the flow of knowledge. The outcomes of the statistical analysis showed that only in some models other variables played a role of importance (sizedif, agedif, vcdif).

Does proximity work this way?

As seen in the theoretical framework a vast amount of theory exists on the beneficial relation between proximity and knowledge flow. One of the quotes that best describes this relation is the following excerpt from Whittington et al. (2009):

> "Knowledge flows across organizational boundaries in all industries, but the intensity and effects of such streams are heightened by spatial proximity (Jaffe, 1989; Gertler, 1995; Maskell and Malmberg, 1999, see Whittington et al. 2009 p92).

As seen from the theoretical framework physical proximity is a facilitating factor for the spreading of knowledge. However this study makes clear that the microphysical effect is negligible on the Leiden Bioscience Park. There are several considerations to this conclusion. For instance, in this study the business is taken as starting point and it is assumed that the location of these businesses on the park might be of importance. Of course, the business itself does not in any way communicate with another business, this is done by its employees. The underlying observation is thus that knowledge flow between (employees of different) organizations is not related to their exact location on the park. Assumed that there is knowledge flow (and there is), microphysical proximity is not a substantial factor in it. In order to study the drivers for the flow of knowledge, other factors should be investigated.

One of the questions that arises from the finding that microphysical proximity is not a significant factor is whether permanent co-location is needed for extracting the benefits of being together? An even more profound question is whether the science park in itself is needed? There are of course benefits of being together, but as we have seen proximity is relative. Do we still need the rather strict boundaries prescribed by the local government? Would the science park not work as well when the companies scatter around the city of Leiden? These questions cannot be answered from the current study, however they are induced by the findings that when microphysical proximity does not have an influence, the boundaries of 'proximity' could be stretched considerably.

If proximity did not work this way, why co-locate?

Due to the negligible effect of microphysical proximity with respect to knowledge flow, the question arises why to co-locate in the first place. Apart from the basic economic advantages, we should not underestimate the effect of the deliberate construction of the bioscience park. As indicated the bioscience park is a designated area within the city of Leiden. Only bioscience or biopharmaceutical companies are allowed to settle on the park. Moreover, companies on the park benefit from several favourable (tax) conditions. Creating the cluster is thus done deliberately. Reasons are of course to ensure that the graduates from the Science faculties of the Leiden University stay within the proximity of Leiden and to attract business that might otherwise have settled in another city. The clustering of bioscience companies is thus also a form of city marketing for the municipality of Leiden. Also for companies present on the bioscience park there are positive effects. Group identification, everyone walking on the park has some involvement with bioscience, contributes to this. Moreover, start-ups might benefit from the positive name that the bioscience cluster has. Identification with the group (cluster), its quality and well known discoveries thus has a positive effect. This is similar to the well-known example of companies having a mailbox on the university campus of Cambridge in order to co-benefit from the positive name this city has due to its excellent scientific reputation.

What if microphysical proximity does have an effect, only the current layout of the LBP does not support it?

It should be kept in mind that this study only looked at the effects of microphysical proximity on knowledge flow <u>on the Leiden Bioscience Park</u>. While ascertaining that the effect of microphysical proximity is negligible we should not falsely extrapolate these findings to other (bio)science parks in the Netherlands, let alone in the rest of the world. There is a chance that due to specific characteristics of the LBP the flow of knowledge is poorly supported. In other words, we have assumed a homogeneous area in which *every meter is a meter.* However, it might wel be the case that due to the urban plan, which is not an orthogonal grid with no holes, outcomes are more subtle. Figure 8.1 gives a graphical representation of this concern. On the left side the assumption of the homogenous area is portrayed, while on the right side an interpretation of the area is given. The area is far from homogenous, with open areas (sports field, park), this might have a rather high influence on the outcomes of this study.

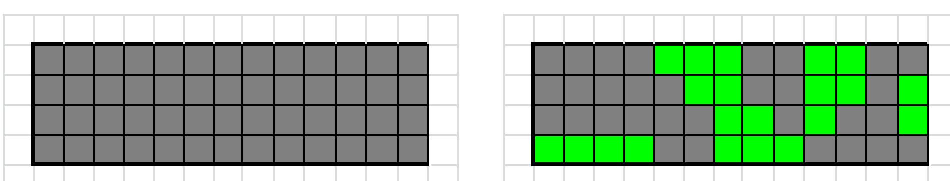

Figure 8.1 Homogenous area

Could we be understating the effects?

There is a chance that the effects are understated within this model. In the theoretical framework the concept of absorptive capacity is introduced. With this concept is meant the extent to which companies can understand each other. They should *speak the same language* in order to communicate with each other. In this study the absorptive capacity is assumed equal for every relation. We are thus pretending that every company can *understand* any other company. 1891 possible cooperative relations were assumed. However it might well be the case that this assumed compatibility is not a valid representation of real life. Further research is needed to assess what factors play a role in the establishment of a knowledge flow relation between two companies. It could well be that what is perceived in this study as relatively few cases is in fact a satisfying score. Figure 8.2 shows this concern; on the left is portrayed that it is assumed that every yellow company could have a knowledge flow relation with each other. Whereas on the right it is indicated that only one company [!] could have engaged in knowledge flow.

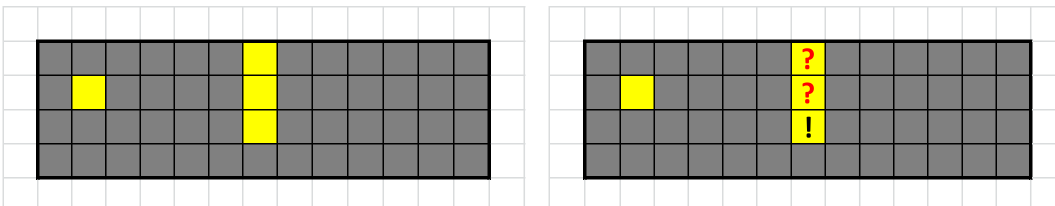

Figure 8.2 Absorptive capacity

Is the use of Binary Logistic Regression justified?

Although the use of binary logistic regression has some considerable advantages over the use of an ordinary least square analysis (as described in the theoretical framework) There are however some considerations that should be kept in mind.

The analysis assumes independent observations. This means that there is no relation between any two observations. In some cases this assumption is violated. Primarily this is the case for the Leiden University and the Leiden University Medical Centre. These large institutions have multiple patents, several R&D cooperations with others and many of their employees have jobhopped. The extent to which this influences the outcomes of the analysis is hard to predict. However it should be kept in mind.

Another concern regarding the use of BLR in this study is that the creation of the model is data-driven. The availability and accessibility of data has undoubtedly steered the form in which this study was performed. Since Science Based Business kept track of several variables it was easy to incorporate those in the study, whereas information that is hard to find (such as profit-and-loss statements, growth indicators and the like) were not taken into account. These variables could however contain valuable information on the performance and success of companies in the (growth as a measure of success, profit and loss statements.

A complicating factor in the interpretation of the results is the persistent negative value for the constant. Although it is not uncommon to have a negative constant (see for instance Hailpern and Visintainer 2003), in this study it is rather hard to explain why. Most logical explanation would be that there are considerable barriers to knowledge flow. Knowledge flow thus only seldomly occurs in very specific cases.

Has knowledge flow been operationalized correctly?

In a way what is called 'tacit knowledge' in this study is already codified knowledge, especially when it comes to patents. Patents are a clear end product of a process in which tacit knowledge has been used. However not every flow of tacit knowledge leads to a patent, on the contrary. The co-patent production was limited to only eight on the park. Presumably the same counts for R&D cooperation and labour mobility. Only after tacit knowledge has spread the two interactors know from each other whether a cooperation of jobhop would be beneficial. The spread of tacit knowledge not necessarily leads to R&D cooperation; both actors could as easily engage in a normal buyer-supplier relation and in this way extract the benefits of each other's knowledge. This tacit knowledge spread is not taken into account in this study. Another consideration is that while there are nearly twelve thousand employees working on the Leiden Bioscience Park only a fraction of those were incorporated in the job-hop variable. This is in itself not troublesome, since the approach of only following CEO's and CSO's is used more often. However as we have seen in the theoretical framework the dispersion of tacit knowledge does not only run via the captains of industry, also via the incremental knowledge build-up in low tech areas on the park (Laestadius 1996; Maskell 1998). In this study this is completely left aside and we run the risk of completely overseeing this type of tacit knowledge flow.

Has proximity been operationalized correctly?

The way in which microphysical proximity is measured could be improved. In this study microphysical proximity is treated as inverted distance, which is in itself is not incorrect. However it is a rather basic form of measuring the proximity between two places. The business is taken as scale level, whereas it would have been better to take *the employee* as starting point. Measuring traveling distance (unobstructed walkways), measuring in time (effort of getting somewhere else), measuring costs (physical or psychological barriers) can improve the approximation of real life situation. The question that arises, however, is whether a more sophisticated approach would have led to different conclusions.

Was the inclusion of other factors justified?

Besides the variables microphysical proximity and knowledge flow, there were other variables included in the model. It was hypothesized that these variables could have an effect on knowledge flow and thus were added as covariates. These were size difference, age difference and value chain difference. As seen in the results and conclusions section these factors are of importance.

Value chain difference

The relation between value chain difference and knowledge flow is considered significant and positive only for one chain, which leads to the conclusion that knowledge flows vertically. As indicated however Maskell (2000) would argue that hopping vertically means that there is increased level of cooperation instead of competition (Hypothesis 2). However what is remarkable in this respect is the strong and positive relation between labour mobility and R&D cooperation, as will be discussed in the following section. Competitive action such as luring away each others employees should not coincide with a very cooperative action such as R&D cooperation and joint venture establishment.

Labour mobility and R&D cooperation

One of the outcomes of the regression analysis is a positive, strong and significant relation between labour mobility and R&D cooperation. In the Conclusions section it was stated that from the data it is obvious that the relation runs in both directions. However the picture might be somewhat more nuanced. What actually should be concluded is that R&D cooperation and labour mobility coincide and co-occur . Since we do not know from the data whether the one precedes the other it is hard to tell how the relation actually functions. One could easily argue that once an employee has switched from company he is more likely to observe and discover the benefits of a potential relation. He might find out that the presumed competitors might in fact have complementary knowledge that could lead to a mutually beneficial relation. A competitive switch thus might lead to a cooperative action. On the other hand it could well be that two companies engaging in a cooperative action learn to know each others employees, and the employees learn to know the other company and the employment benefits of the potential new employer. A cooperative action might thus lead to a competitive action. Is one of the two relations more or less likely to occur? The initial and most logical reaction might be to say that the latter is less likely to be found in practice, since a competitive switch of employees puts an immediate pressure on the cooperative relationship. And the first relation is also in line with some findings of entrepreneurial behaviour (QUOTE XXX) where it is argued that the in-depth knowledge of a subject is needed to engage in entrepreneurial activities, such as starting a R&D cooperation.

What about the factor time?

Another drawback is that time is not taken into account. As is clear from several studies (for instance Neffke et al. 2010) business parks evolve over time. One can distinguish between several development phases and in each development phase the park has its own characteristics. As Neffke et al. (2010) conclude that ‘with increasing levels of maturity, industries experience rising intra industry spillovers, but declining inter-industry spillovers’ (p63). Further research on this subject, specified for the Leiden Bioscience Park, could improve the understanding of knowledge flow within the park. Knowing in which development phase the park and its businesses are is of importance to stimulate inter or intra industrial contact.

Overall conclusion (take home message)

This study shows that microphysical proximity is of negligible influence on the flow of knowledge. However this does not mean that the existence of the Leiden Bioscience Park is not justified. Other characteristics of the businesses are of importance, as well as intangible factors such as group cohesion and city marketing effects.

9. Recommendations

- Although the Binary Logistic Regression does portray clear outcomes, it is advisable to compare these findings with other research methods. As indicated there is little quantitative information available on the subject of knowledge flow. Using the constructed database in another way might lead to even more insightful outcomes.
- Since knowledge is one of the production factors to businesses in the bioscience industry, knowing what factors influence the creation and/or spread of knowledge is of vital commercial importance. Having established that microphysical proximity is of negligible influence on the Leiden Bioscience Park, the focus should be moved to other factors that could improve the flow of knowledge. Perhaps that human encounters on congresses, informal meetings, and post-doc events might have considerable influence on the likelihood that knowledge is spread. Performing a study similar to the study of the veldacademie might be insightful. In this study citizens of Rotterdam city center were GPS-tracked in order to study the patterns in their whereabouts (see the website of both www.veldacademie.nl and bk.tudelft.nl). Knowing where employees of the bioscience park go, stay and meet each other is presumably very insightful.
- Placing this study in a larger perspective by performing a similar study on other bioscience parks might lead to interesting results. Conclusions could be in line with the findings from this study, e.g. the scale level of microphysical proximity does not play a significant role. Or conclusions could differ e.g. on other parks there is a relation between microphysical proximity and knowledge flow. If the latter is the case, this could open a new era of attention to the field of microphysical proximity, since then specific characteristics of the parks should be compared. As indicated in the discussion it could be that the current layout of the Leiden Bioscience Park does not support knowledge flow.
- Another recommendation might be to look into the specific cases in which knowledge flow has been observed. A study that performs an in depth analysis on how these companies got in contact with each other might provide valuable insights. It might well be that a shared background (graduated from the same study, membership of the same fraternity/sorority, having lived in the same dormitory) is of large influence.
- Since the process of innovation and knowledge creation deals with tacit knowledge, a way should be found to trace this typical form of knowledge. In this study the end products of tacit knowledge flow have been taken as dependent variable. However a way should be sought to detect the spread of tacit knowledge earlier in the process of knowledge creation.
- Since there are indications that the bioscience park is a rather tightly knit network (a large number of employees studied in Delft, university playing a dominant role etc.) it could be very interesting to follow the activities of the chief-level on social media (twitter, linked-in, etc.). Getting to know the characteristics of the network might contribute considerably to understanding how a business park functions and what network factors play a role.

10. Reference List

Audretsch, D.B. 1998 Agglomeration and the location of innovative activity. Oxford Review Economic Policy, Issue 14: pp.19–28

Audretsch, D.B. 2003 Innovation and spatial externalities. International Regional Science Review, Issue 26: pp.167–174

Audretsch, D.B., Feldman, M.P. 2004 Knowledge spillovers and the geography of innovation. In: Henderson V, Thisse JF (eds)Handbook of regional and urban economics, Vol 4. Elsevier,Amsterdam, pp.2713–2739

Breschi, S. and Malerba, F. 2005 Clusters, Networks and Innovation. Oxford University Press.

Boschma, R., Eriksson, R., Lindgren, U. 2009. How does labour mobility affect the performance of plants? The importance of relatedness and geographical proximity Journal of Economic Geography , Issue 9: pp. 169–190

Boschma, R. (2005) 'Proximity and Innovation: A Critical Assessment', Regional Studies, Vol.39, Issue 1: pp.61 – 74

Burger, M.J., Oort van, F.G., Frenken, K., Knaap van der, B. 2009 Networks and economic agglomerations: introduction to the special issue. Tijdschrift voor Economische en Sociale Geografie, Vol. 100, Issue 2: pp.139–144

Capello, R. 2009 Indivisibilities, synergy and proximity: the need for an integrated approach to agglomeration economies. Tijdschrift voor Economische en Sociale Geografie, Vol. 100, Issue 2: pp. 145–159

CONNECT 2010 Interview with Nettie Buitelaar. Issue 33. 25th March 2010. retrieved via http://www.indodutchconnect.com/articles/leiden-bioscience-park_84.html

Cruz, S.C.S., Teixeira, A.A.C. 2009 The Evolution of the Cluster Literature: Shedding Light on the Regional Studies-Regional Science Debate. Regional Studies, Vol. 44, Issue 9: pp. 1263–1288

Freeman, C. 1987 Technology policy and economic performance: Lessons from Japan, London: Frances Pinter.

Gertler, M. S. 1995 Being there: Proximity,organization, and culture in the development and adoption of advanced manufacturing technologies. Economic Geography, Issue71: pp. 1–26

Hailpern, S.M., Visintainer P.F. 2003 Odds ratios and logistic regression: further examples of their use and interpretation. The State Journal, Vol. 3, Issue 3: pp.213-225

Hervas-Oliver, J.L., Albors-Garrigos J. 2009 The Role of the Firm's Internal and Relational Capabilities in Clusters: When Distance and Embeddedness are Not Enough to Explain Innovation. Journal of Economic Geography, Vol. 9, Issue 2: pp.263-283

Howells, J. 1996 Tacit knowledge, innovation and technology transfer. Technology Analysis and Strategic Management, Issue 8: pp.91-106

Jaffe, A.B. 1989 Real effects of academic research. American Economic Review, Issue 79: pp.957–970

Jaffe, A. B., Trajtenberg, M., Henderson, R. 1993 "Geographic localization of knowledge spillovers as evidenced by patent citations." Quarterly Journal of Economics, Vol. 79: pp957–970

Jenkins, M., Tallman, S. 2010. The shifting geography of competitive advantage: clusters, networks and firms. Journal of Economic Geography 10 pp. 599–618

Laestadius 1998 The relevance of science and technology indicators: the case of pulp and paper. Research Policy, Vol. 27, Issue 4, pp.385-395

Marshall, A. Principles of Economics. retrievable from www.econlib.org/library/Marshall/marP.html

Maskell, P., Malmberg, A. 1999 Localized learning and industrial competitiveness. Cambridge Journal of Economics, Issue 23: pp.167–186

MIBAC (Ministero per I Bene e le Attativa Culturali) 2008 Nuovi Uffizi retrieved via www.nuoviuffizi.it/english/complesso/fabbriche/ on 29-05-2011

Moodyson, J., Jonsson, O. 2007 Knowledge collaboration and proximity: the spatial organization of biotech innovation projects. European Urban and Regional Studies, Vol.14, Issue 2: pp.115–131

Morgan, K. 2004 The exaggerated death of geography: learning, proximity and territorial innovation systems. Journal of Economic Geography, Volume 4, Issue 1: pp.3-21

Neffke, F., Henning, M., Boschma, R., Lundquist, K.J., Olander,L.O. 2010 The Dynamics of Agglomeration Externalities along the Life Cycle of Industries. Regional Studies, Vol. 45: pp49-65

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

Sorenson, O. 2005 Social Networks and the Persistence of Clusters: Evidence from the Computer Workstation Industry, pp.297-316. In Breschi, S. and Malerba, F. (eds.) 2005 Clusters, Networks and Innovation. Oxford University Press.

Vitiello, D. 2005 De/Industrialization and Sub/Urbanization: New Perspectives on Industrial Place-Making in North America. Journal of Planning History, Issue 4: pp254-267

Whittington, K.B., Owen-Smith, J., Powell, W.W. 2009 Networks, Propinquity, and Innovation in Knowledge-intensive Industries. Administrative Science Quarterly, Vol. 54: pp 90–122

Appendix 1 – SPSS output

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	526.657[a]	.037	.138

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4.949	8	.763

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	SIZEDIF	.000	.000	16.766	1	.000	1.000
	AGEDIF	.066	.018	13.362	1	.000	1.068
	VCDIF			12.074	5	.034	
	VCDIF(1)	1.100	.552	3.973	1	.046	3.004
	VCDIF(2)	-.297	.735	.163	1	.686	.743
	VCDIF(3)	.310	.744	.174	1	.677	1.363
	VCDIF(4)	.533	.807	.437	1	.509	1.705
	VCDIF(5)	-16.914	2634.457	.000	1	.995	.000
	DISTANCE	-.001	.000	4.333	1	.037	.999
	Constant	-4.375	.575	57.904	1	.000	.013

a. Variable(s) entered on step 1: DISTANCE.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	87.965[a]	.014	.235

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5.587	8	.693

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	SIZEDIF	.001	.000	22.344	1	.000	1.001
	AGEDIF	-.115	.061	3.542	1	.060	.892
	VCDIF			.058	5	1.000	
	VCDIF(1)	16.858	2699.174	.000	1	.995	2.1E+07
	VCDIF(2)	.708	3574.104	.000	1	1.000	2.030
	VCDIF(3)	16.575	2699.174	.000	1	.995	1.6E+07
	VCDIF(4)	.998	4745.040	.000	1	1.000	2.713
	VCDIF(5)	.884	3755.418	.000	1	1.000	2.421
	DISTANCE	-.001	.001	.369	1	.543	.999
	Constant	-21.199	2699.174	.000	1	.994	.000

a. Variable(s) entered on step 1: DISTANCE.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	97.281[a]	.009	.152

a. Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	8.815	8	.358

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	SIZEDIF	.001	.000	23.560	1	.000	1.001
	AGEDIF	-.129	.065	3.906	1	.048	.879
	DISTANCE	-.001	.001	.547	1	.460	.999
	Constant	-4.787	.592	65.300	1	.000	.008

a. Variable(s) entered on step 1: DISTANCE.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	95.105[a]	.010	.172

a. Estimation terminated at iteration number 9 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	10.283	8	.246

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	AGEDIF	-.145	.068	4.453	1	.035	.865
	SIZEDIF	.001	.000	20.528	1	.000	1.001
	DISTANCE	-.001	.001	.557	1	.456	.999
	LABMOB	.653	1.358	.231	1	.631	1.921
	RDCOOP	1.735	1.309	1.757	1	.185	5.669
	Constant	-4.787	.587	66.492	1	.000	.008

a. Variable(s) entered on step 1: LABMOB, RDCOOP.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	305.311[a]	.027	.156

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7.168	8	.519

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	SIZEDIF	.000	.000	8.464	1	.004	1.000
	AGEDIF	.090	.024	13.901	1	.000	1.094
	VCDIF			6.954	5	.224	
	VCDIF(1)	1.742	1.052	2.743	1	.098	5.709
	VCDIF(2)	.721	1.186	.370	1	.543	2.057
	VCDIF(3)	.835	1.264	.437	1	.509	2.306
	VCDIF(4)	.691	1.462	.223	1	.637	1.995
	VCDIF(5)	-15.426	2577.235	.000	1	.995	.000
	DISTANCE	-.002	.001	8.924	1	.003	.998
	Constant	-5.619	1.067	27.712	1	.000	.004

a. Variable(s) entered on step 1: DISTANCE.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	319.958[a]	.019	.111

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	6.735	8	.565

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	SIZEDIF	.000	.000	8.141	1	.004	1.000
	AGEDIF	.099	.024	16.294	1	.000	1.104
	DISTANCE	-.001	.000	9.582	1	.002	.999
	Constant	-4.393	.376	136.449	1	.000	.012

a. Variable(s) entered on step 1: DISTANCE.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	282.385[a]	.038	.224

a. Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	11.830	8	.159

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	AGEDIF	.079	.026	9.194	1	.002	1.082
	SIZEDIF	.000	.000	5.272	1	.022	1.000
	DISTANCE	-.002	.001	10.699	1	.001	.998
	RDCOOP	3.112	.458	46.118	1	.000	22.455
	PATENT	1.276	1.229	1.078	1	.299	3.583
	Constant	-4.399	.385	130.917	1	.000	.012

a. Variable(s) entered on step 1: RDCOOP, PATENT.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.288	.000
	Cramer's V	.288	.000
N of Valid Cases		1891	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		1.000[a]
N of Valid Cases	1891	

a. Binomial distribution used.

LABMOB * RDCOOP Crosstabulation

Count

		RDCOOP		Total
		.00	1.00	
LABMOB	.00	1829	26	1855
	1.00	25	11	36
Total		1854	37	1891

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	323.496[a]	.021	.122

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	6.673	8	.572

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	SIZEDIF	.000	.000	1.969	1	.161	1.000
	AGEDIF	.095	.024	15.339	1	.000	1.099
	VCDIF			6.154	5	.291	
	VCDIF(1)	.542	.635	.729	1	.393	1.720
	VCDIF(2)	-.831	.934	.793	1	.373	.435
	VCDIF(3)	-1.044	1.181	.781	1	.377	.352
	VCDIF(4)	.095	.954	.010	1	.921	1.100
	VCDIF(5)	-16.853	2629.264	.000	1	.995	.000
	DISTANCE	.000	.000	.116	1	.733	1.000
	Constant	-5.153	.691	55.615	1	.000	.006

a. Variable(s) entered on step 1: DISTANCE.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	341.345[a]	.012	.069

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5.775	8	.672

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	AGEDIF	.111	.023	23.406	1	.000	1.117
	DISTANCE	.000	.000	.376	1	.540	1.000
	Constant	-5.175	.408	160.627	1	.000	.006

a. Variable(s) entered on step 1: DISTANCE.

MODEL	DV	IV
1	KNOWFLOW	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.1	PATENT	SIZEDIF, AGEDIF, VCDIF, DISTANCE
2.2	PATENT	SIZEDIF, AGEDIF, DISTANCE
2.3	PATENT	SIZEDIF, AGEDIF, DISTANCE, LABMOB, RDCOOP
3.1	LABMOB	SIZEDIF, AGEDIF, VCDIF, DISTANCE
3.2	LABMOB	SIZEDIF, AGEDIF, DISTANCE
3.3	LABMOB	SIZEDIF, AGEDIF, DISTANCE, PATENT, RDCOOP
4.1	RDCOOP	SIZEDIF, AGEDIF, VCDIF, DISTANCE
4.2	RDCOOP	AGEDIF, DISTANCE
4.3	RDCOOP	AGEDIF, DISTANCE, PATENT, LABMOB

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	301.174[a]	.033	.188

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	8.898	8	.351

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	AGEDIF	.095	.025	14.168	1	.000	1.099
	DISTANCE	.000	.000	1.695	1	.193	1.000
	PATENT	1.898	1.165	2.652	1	.103	6.671
	LABMOB	3.118	.442	49.728	1	.000	22.607
	Constant	-5.498	.440	156.102	1	.000	.004

a. Variable(s) entered on step 1: PATENT, LABMOB.

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.288	.000
	Cramer's V	.288	.000
N of Valid Cases		1891	

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Chi-Square Tests

	Value	Exact Sig. (2-sided)
McNemar Test		1.000[a]
N of Valid Cases	1891	

a. Binomial distribution used.

LABMOB * RDCOOP Crosstabulation

Count

		RDCOOP		Total
		.00	1.00	
LABMOB	.00	1829	26	1855
	1.00	25	11	36
Total		1854	37	1891