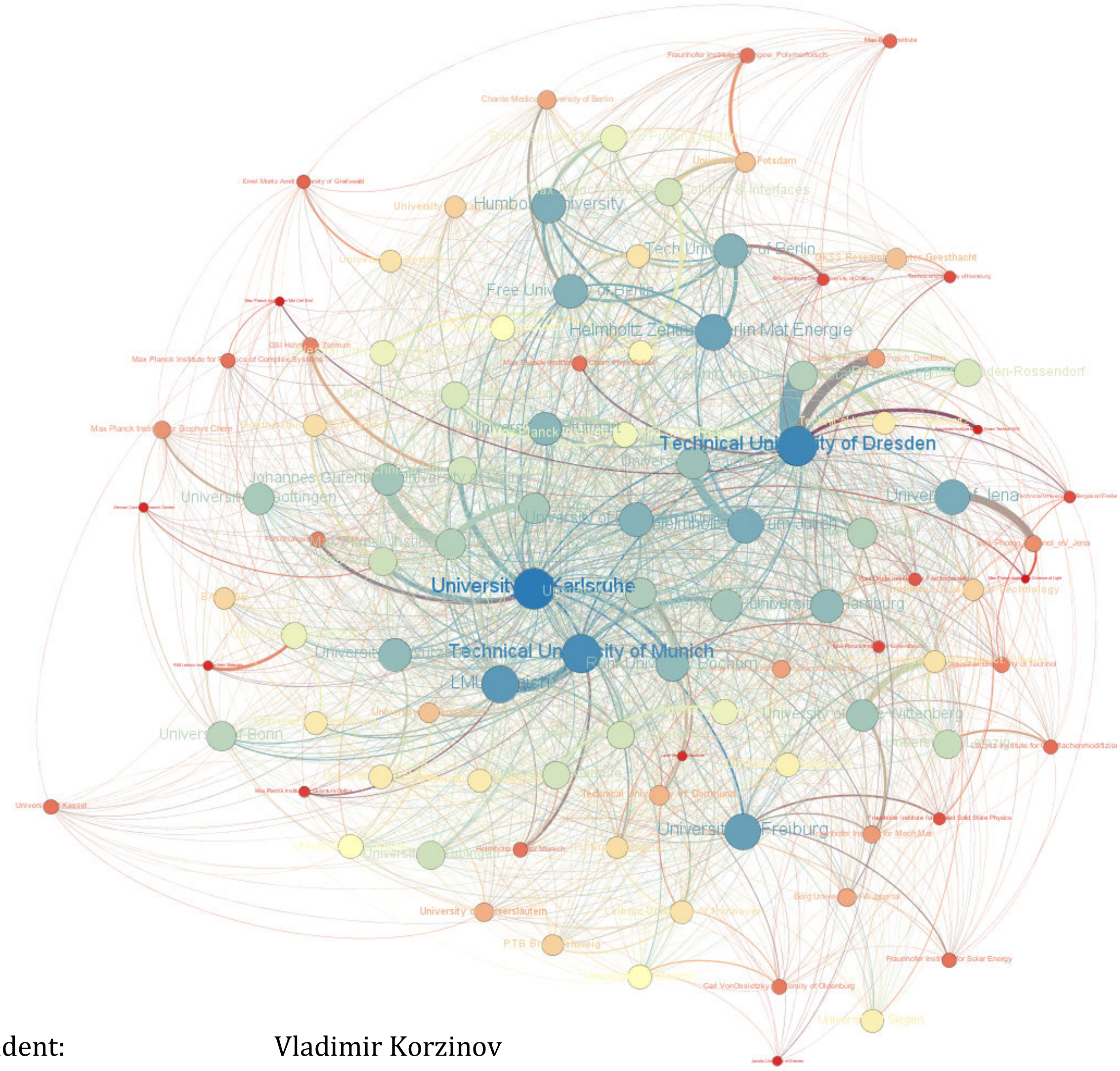
DELFT UNIVERSITY OF TECHNOLOGY

Master Thesis

Proximity and Collaboration in German Nanotechnology Network

Faculty of Technology, Policy and Management
Master programme Management of Technology



Student: Vladimir Korzinov

Full professor (Chair) Alfred Kleinknecht
1st Supervisor Claudia Werker
2nd Supervisor Scott Cunningham

Acknowledgements

This master thesis concerns the analysis of the relationships between different dimensions of proximity and collaboration in German nanotechnology network. It is the last assessment for my master study Management of Technology at the TU Delft.

My engineering diploma obtained in Russia was related to nanotechnology. While looking for my master thesis topic I always wanted to write something related to my background. When I saw a thesis proposal from Claudia and Scott to study the nanotechnology networks it was a perfect match. I suggested the case of Germany because I always had an eye on this country and wanted to learn its language. Claudia asked Alfred to be my professor because of his expertise in economics and innovation.

I would like to thank my family and flat mates for their endless support that helped me to stay focused and motivated. I am so thankful to my supervisors for the opportunity to write my final work at TU Delft relating to my bachelor background. It gave me additional inspiration. I would like to express my appreciation to Alfred for his relaxed support. Without various discussions with Scott I won’t be able to equip myself with all the necessary tools and instruments to conduct this research. I would like to thank Claudia for her theoretical support, consultations about German research system and help in writing this report.

Abstract

The processes of the establishment and facilitation of inter-organizational collaboration in German nanotechnology network are differently influenced by proximity dimentions. Collaborations are very important for newly emerging technologies such as nanotechnology. With this master thesis we focus on howgeographical, organizational and technological proximities influence collaboration activities in nanotechnology. We base the analyses on publication data for the last three years. We were able to show which of these dimensions play a role in establishing and facilitatingcollaborations using methods of regression analysis. While geographical and technological proximity directlyaffect both of these processes, organizational is only affecting the establishment of collaborations. We explain why proximity and collaboration are related in such ways. Based on that we offer management and policy recommendations.

Table of Contents

Acknowledgements 2
Abstract 3
Table of Contents 4
Figures and Tables 6
1. Introduction 7
Research question 9
2. The development of the conceptual model 10
Literature review 10
The concept of collaboration in relation to knowledge processes 11
The concept of proximity in relation to knowledge processes 12
The concept of proximity in relation to collaboration 13
Conclusion of the literature review 14
Conceptual model 14
Collaboration 14
Proximity dimensions 16
3. Research methodology 21
Research strategy 21
Data collection methods 21
Population of the study 22
Sample of the study 23
Results of the sampling procedure 24
Research variables 27
Physical proximity variable 27
Systemic proximity variables 27
Organizational proximity variables 28
Technological proximity variable 28
Interactions between research variables 29
Descriptive statistics of research variables 30
Multicollinearity 31
4. The influence of proximity on the likelihood of collaboration 32
Dependent variable 32
Model equation 32

Specifications of the application of the first model 33
Regression results of the first model 33
5. The influence of proximity on the intensity of collaboration 36
Dependent variable........ 36
Model equation........ 36
Specifications of the application of the second model........ 36
Regression results of the second model 37
6. German nanotechnology network as a part of European network........ 40
Dependent variable........ 40
Model equation........ 40
Specifications of the application of the third model 40
Regression results of the third model........ 41
7. Limitations and assumptions 44
8. Conclusion 45
9. Management and policy recommendations........ 48
10. Appendix A 49
11. Appendix B 51
12. Appendix C 52
13. Appendix D 55
14. Appendix E 59
15. Bibliography 67

Figures and Tables

Figure 1 - Research framework 10
Figure 2 - Proximity dimensions relevant for inter-organizational collaboration (J. Knoben, 2006)..... 16
Figure 3 - The conceptual model 20
Figure 4 – Calculation of publication productivity of organizations 24
Figure 5 - German Nanotechnology Organizations 25
Figure 6 – Constructed Network of 100 German Organizations 26
Figure 7 – Collaborations intensity (log scale) depending on the technological proximity 43

Table 1 - Spelling variations of a university name in the Web of Knowledge database 24
Table 2 - Top 10 German nanotechnology organizations 25
Table 3 - Dummy coding of organizational variables 28
Table 4 - Number of publications in six major nanotechnology science categories 29
Table 5 - Independent regression variables and their interactions 30
Table 6 - Systemic proximity in German nanotechnology network 31
Table 7 - Variance Inflation Factors of research variables 31
Table 8 - The prediction improvement of the logistic regression model 33
Table 9 - Binary logistics regression coefficients 34
Table 10 – Multiple linear regression summary 37
Table 11 - ANOVA test of the multiple linear regression 37
Table 12 – Multiple linear regression coefficients 38
Table 13 - Negative binomial regression coefficients 42

1. Introduction

Nowadays nanoscience and technology (NST) is driving innovation and technological change by breaking through the boundaries of existing technologies. The European Commission identifies nanomaterials as a key enabling technology(European Comission, 2012). Forecasts exhibit that the global value of products based on NST will grow from €200 billion in 2009 to €2 trillion by 2015(The European Commission, 2012). In the first decade of the 21st century the global investment in nanoscience by governments accounted for more than $65 billions and is far from reaching its limit(Cientifica Ltd., 2011). Most scholars define nanoscience and technology as *"an activity with the investigation of bottom-up and top-down structural arrangements at a physical size below 100 nanometers, where the properties of materials, systems and devices differ significantly from those at a larger scale"*(Kostoff & et.al., 2007).NST already underpins the vast amount of products in various industrial sectors including crucial health, food, environment, energy and transport applications and its share continues to increase. Pushing forward the NST research frontier will help to advance beyond industrial innovation and increase economic prosperity.

Four characteristics distinguish NST from other scientific and technological fields(Salerno & et.al., 2008). NST isinterdisciplinar (1), pervasive (2), at an early stage of development (3) and spread out throughout the world (4).First two peculiarities determine the nature of NST. The interdisciplinary character (1) of nanotechnology follows from the fact that world consists of molecules and atoms. There is no difference between titanium particles or proteins when operating at such a small level. Nanotechnology unites several fundamental sciences to look at the world with the resolution of one billionth of a meter.Due to such a small scale of manipulation the products produced with the help of nanotechnology find their potential applications in many different fields. Being very pervasive (2) makes NST a potential general purpose technology of our time like ICT used to be at the end of 20th century.The third and the fourth peculiarity characterize the dynamic and scale of nanotechnology expansion. First, being on an early stage is a typical characteristic of rapidly developing technological regimes. Second, research breakthroughs in NST are so widely distributed and diverse in terms of their geography, industry, application and market niche that no single organization has enough internal resources to achieve success.The locus of innovation in NSTwas found to be in networks rather than in individual firms because the knowledge base is complex and expanding and the sources of expertise are widely dispersed(Smith-Doerr & et.al., 1996).

Due to these peculiarities, networks of researchers are highly important for nanotechnology.The network structure of nanoscience imposes various challenges. A high degree of both institutional and disciplinary diversity may create problemsfor the management and coordination of such systems.Public and private organizations need to be brought together to carry out collaborative research and development. This process may create communication problems. In the light of this, international and inter-institutional collaboration becomes an imperative for innovation as a consequence of widespread dispersion of knowledge in nanotechnology networks(Pandza & et.al., 2011).The costs and quality of physical products are typically a function of how well the firm's network collaborates in its development and production. Similar, the quality and relevance of research in nanotechnology is dependent on the productivity of the research network and is therefore a function of the collaboration in a network of firms, universities and research institutions.

Collaboration represents a major exchange of knowledge as well as production of new knowledge(Katz & Martin, 1997). The ability of a scholar to effectively communicate with colleagues and to address a broad work spectrum is determining for the process of creation of new ideas(Heinze & Bauer, 2007). Various researchers have realized that inter-organizational learning through collaboration is critical for competitive success of an organization(Dyer & Nobeoka, 2000). The level of collaboration continues to increase in the scientific community. This process is fuelled by decreasing costs of travel and communication, development of ICT and raising importance of interdisciplinary fields like nanotechnology(Katz & Martin, 1997).Research collaboration may take a variety of different forms, from informal meetings between researchers to official partnerships between firms, institutions and universities. The variety of collaborative forms and their importance for the network structures emphasize its central role for nanoscience and technology.Although knowledge networks and collaborations were in the focus of variety of researchers there is still a scope for adding further scientific value(Brenner & et.al., 2011).Therefore, facilitators and hampers of collaboration as well as its advantages and disadvantages represent essential questions for the research.

One of the indispensablefacilitators and sometimes also hampers of collaboration is proximity. Various dimensions of proximity differentlyaffect inter-organizational collaboration. A failure to understand these effects may lead to the wrong management and policy decisions. Such decisions may cause a downfall in communication and cooperation between researchers and scholars which will negatively affect the development of nanoscience.In the following research we shed the light on the influence of proximity dimensions on collaboration in German network of nanotechnology organizations. Knowledge about how different factors influence collaboration can be used by governments, universities and other research institutions as a basis for the management and policy decisions.

We choose Germany because it is one of the locomotives in nanotechnology development in Europe. European Union sustains its leading positions in the NST measured by publications and patent analysis(Youtie, Shapira, & Porter, 2008). Currently 40 per cent of the world publications in nanoscience come from Europe and the vast majority of them as well as patent filing comes fromGermany(Heinze, 2004).The country has been one of the world leaders in innovative technologies for the past few decades. The evolution of German nanotechnology started in early 1990s. The volume of funding increased more than tenfold together with the shift in strategic orientation of investments(Zweck & et.al., 2008). NST is identified as one of the key technologies to promote a *HIGH-TECHSTRATEGY 2020*, driving the creation of new products, procedures and services.The country created an example of a successful nanotechnology network. However, government authorities recognize that there is still unrealized potential in the transfer of knowledge(Federal Ministry of Education and Research, 2012). The knowledge gained in this master thesis may help to find more efficient allocation of collaborative links.

Research question

Following from above the main question of this master thesis is formulated as follows.

How do different dimensions ofproximity affect collaboration and knowledge transfer in the German nanotechnology network?

The remainder of the thesis is structured as follows. Chapter 2 begins with the introduction of necessary knowledge about collaboration and proximity obtained in scientific literature. Afterwards we provide all definitions of all concepts and formulate the hypotheses to be tested. We continue with Chapter 3 elaborating on the research methodology. It begins with the description of the research population and a sample.The paper continues with presentation of the operationalisation of the concepts and a full list of independent variables used in the research. At the end of Chapter 3 we discuss the descriptive statistics of obtained variables and show the constructed network of German organizations. Chapter 4 provides the information about initiators of collaboration. It begins with the description of the regression model and finishes with the analysis of its results. Chapter 5 is aimed to present facilitators of collaboration activity. It first introduces the details of the regression model and then shows the analysis of its results. Chapter 6 elaborates on the third constructed model. It presents its description and results. We continue with Chapter 7 pointing out limitations and assumptions of the carried research. Chapter 8provides conclusions and recommendations for the further studies. Finally Chapter9tells about management and policy implications that follow from our analysis.

2. The development of the conceptual model

Chapter 2 forms a conceptual model for the analysis of German nanotechnology network. Firstly, through the literature review we provide necessary theoretical perspectives on nanotechnology, collaboration, proximity and knowledge processes. Secondly, we formulate definitions of the main concepts used in our research. Finally, we provide eight hypotheses to be tested.

Research framework in Figure 1shows that after having studied the relevant literature on collaboration, proximity, and knowledge processes a) the conceptual framework will be constructed b). The network of German nanotechnology will be analysed using the framework developed c).

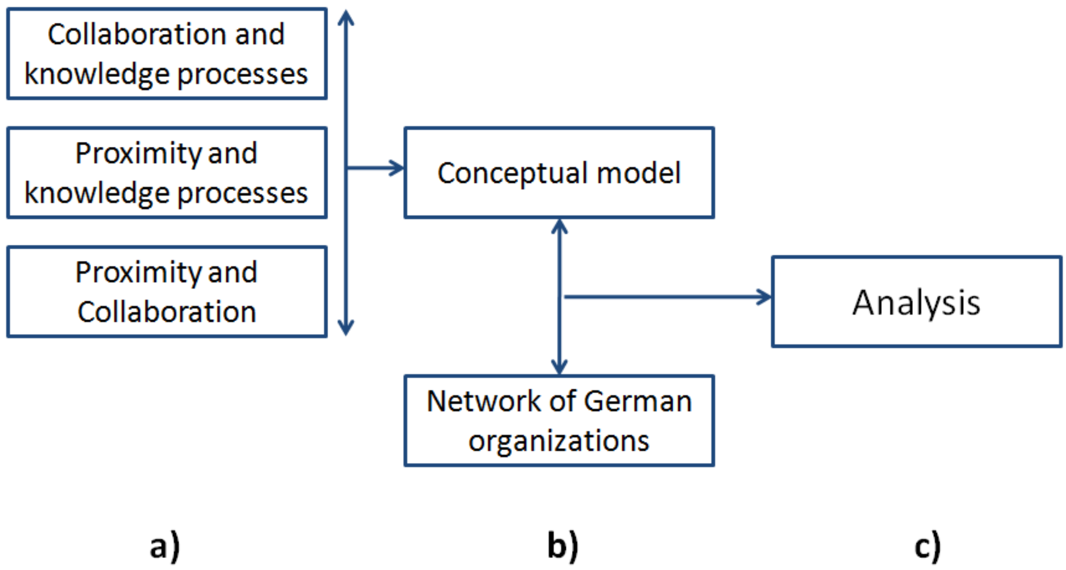


Figure 1 - Research framework

Literature review

Our literature review aims to explain processes happening in nanotechnology networks and to show the importance of proximity and collaboration. We first explore some characteristics of nanotechnology networks and then show the relationships of proximity, collaboration and knowledge processes following from them.

One can see nanoscience and technology as a common denominator for other scientific disciplines. At the nanoscale researchers can easily switch from physics to chemistry and then to biology. The process of nanoscience development is accompanied with the fusion of more traditional scientific disciplines into it. Since the early 1990sresearchers in classical sciences such as chemistry, physics, materials science, and biology have introduced a new dimension of nanoscience(Islam & Miyazaki, 2009).

The interdisciplinarity of NST leads to the creation of avery divergent knowledge. The shift from the old disciplines to emerging ones was followed by a switch in knowledge production from so called “mode 1” to “mode 2” (Jansen & et.al., 2010). “Mode 1” is considered to be disciplinary related and was associated with old established sciences such as physics or chemistry. In contrast, “mode 2” is seen as transdisciplinary knowledge production blurring boundaries between traditional sciences as well as basic and applied research.

Production of a divergent knowledge best benefits from the open research networks with weak ties among groups of scholars (Jansen & et.al., 2010). These network structures fit if a system is pursuing a knowledge exploration strategy(Pandza & et.al., 2011). The strategy implies stimulation of creativity among researchers. Several works showed a positive effect of open networks on the citation scores and scientific performance of researchers (Heinze & Bauer, 2007) (Jansen & et.al., 2010). It was spotted that due to the fusion of other disciplines in NST researchers have to broaden their work spectrum which usually leads to an increased creativity and new ideas(Islam & Miyazaki, 2009)(Heinze & Bauer, 2007).

Apart from the work spectrum the creation of novel ideas also depends on the ability of a researcher to effectively communicate with peers(Heinze & Bauer, 2007).Nanoscienceis usually performed by a small number of large communities or a large number of small communities(Onel & et.al., 2011). This extensive knowledge creation involves not only the discipline diversity but also institutional diversity. In fact nanoscience networks includeindividuals, universities, firms, public policy agents and others(Pandza & et.al., 2011). In these networks parties may benefit from each other's complementary skills while working together. However, the misalignment of objectives inevitably present in diverse networks can hamper effective collaboration.

The collaboration activities in German research system were studied by Heinze and Kuhlmann(Heinze & Kuhlmann, 2008).Authors stress its high level of institutional fragmentation. Focusing on the two subfields of nanoscience and technology, namely nano-electronics and nano-interfaces, authors identify governance structures that support or hinder scientists' efforts to engage into collaboration across institutional boundaries. There are different motives for scientists to cooperate with theircolleagues and other organizations such as an expansion and improvement of their research capacities, benefit from organizational complementarities or enhancement of their visibility within the research field(Heinze & Kuhlmann, 2008).It is widely recognized that research collaboration is a key mechanism for knowledge production and diffusion.

The concept of collaboration in relation to knowledge processes

Scientific collaboration is tightly linked with different knowledge processes. The distinction between them lays between production and application of knowledge.Newly discovered knowledge may be highly ranked by the scientific community, however it may not find a valuable application. And vice versa, a useful application may be seen as outdated from theory perspective comparing to the research frontier. Collaboration activities that promote one process can be tolerant to another.

Lavie and Droripoint out that academic collaboration contributes to the knowledge creation(Lavie & Drori, 2012). A transdisciplinary "mode 2" knowledge production implies a central role of collaboration in innovation. Due to the fusion of disciplines into NST people with different backgrounds cooperate creating and applying new ideas. The process is accompanied by a number of benefits. Firstly, with the increasing number of academic collaborators scientists can complement their skills by relying on the competence of their partners. Secondly, they increase their social capital which facilitates an idea generation and enhances reputation in the scientific community. Thirdly, researchers get access to the tacit knowledge that cannot be documented and distributed. This process may also stimulate creativity. Fourthly, universities share costs and resources by doing projects together. And finally, collaboration enhances the division of labor increasing the efficiency of performed tasks(Katz & Martin, 1997). However, there are certain downsides of this process. As the number of collaborators increases the process may strain the ability to coordinate multiple relationships, integrate knowledge flows and impose constrains on internal learning resulting in a less fruitful knowledge creation(Lavie & Drori, 2012).

Whereas academic collaboration corresponds to the exploration of knowledge, industry collaboration affects its application. Strong and weak linksamong firms may facilitate the exploitation of new knowledge. In order to survive in the world's competition industry sector is always oriented towards the needs of customers. Firms are always motivated to profit from a gained knowledge. University relationships with industry may provide necessary resources for commercialization of a developed technology. Therefore, such cooperation can provide an efficient division of labor when one party is focused on the exploration and the other on the exploitation of knowledge. However, the process may be hampered by the increasing number of partners because of intellectual property rights and spreading the efforts of scientists across too many industry engagements(Lavie & Drori, 2012).

In both processes the availability of internal resources may substitute network collaboration. Resource-poor organizations are most likely to benefit from mutual activities with others(Lavie & Drori, 2012). However, not only internal resources may hamper these activities. The problem of the mutual knowledge may become an issue when maintaining collaboration with a geographically dispersed partners(Cramton, 2001). Establishment of the shared knowledge base is crucial as it increases the chances that parties will understand each other. Moreover, it allows researchers to operate on a more comprehensive level being aware that your partner understands you.

The concept of proximity in relation to knowledge processes

Knowledge processes are not only affected by collaboration intensity but also by the proximity of researchers. Different forms of proximity are distinguished in scientific literature. Researchers may be close to each other geographically, or culturally belonging to the same community, or institutionally sharing the same innovation system and in many other forms. All these dimensions of proximity affect the knowledge produced and distributed across the network.

The value of the proximity to the knowledge processes depends on the type of knowledge. Three main types can be distinguished: simple, complex and moderate complexity(Rivkin & et.al., 2006). In order to better understand the criteria of complexity levels knowledge can be conceptualized as a recipe that one researcher gives to another(Rivkin & et.al., 2006). The more complicated a recipe is and the more rare its ingredients are the higher the level of knowledge complexity is. Simple knowledge can be easily transferred and even distant members of the network will be able to accumulate it. Complex knowledge is hard to grasp even by the proxy members of research community. Only third moderate complexity knowledge is influenced by the proximity.

On the example of US patents from 1990s, Sorenson and Rivkin were able to show that organizational and geographical proximity positively influence the diffusion of moderate complexity knowledge in the network. The problem of mutual knowledge can be also seen here. The more ingredients of the recipe collaborating parties share, the smoother would be the flow of knowledge between them due to the proximity.

Additionally, Boschma discusses the concept of proximity in relation to an interactive learning and innovation. He shows that there is a negative lock-in effect of spatial proximity(Boschma, 2005). Being too close to some parties a researcher may lose the capability of acquiring new knowledge through interactions with more distant actors.Therefore, a balance is required between local and distant links of researchers.

We would like to emphasize that it is impossible to study geographical proximity in isolation but it should be assessed in relation to its other dimensions(Boschma, 2005). Geographical proximity may stimulate a knowledge creation.However,it can do so only in combination with other forms of proximity such as cognitive, institutional, and social among others. Moreover, these forms of proximity may also serve as a substitute for spatial proximity. In fact a transdisciplinary 'mode 2' knowledge production requires different forms of proximity between researchers in order to establish and maintain effective relationships that will lead to a productive output. Additionally it has to be mentioned that for a smooth knowledge transition proximity requires some but not too greatdistance between actors(Boschma, 2005).

We could see that theinterdisciplinarity and the diversity of institutions in nanotechnology networks lead to complex knowledge processes that are affected by proximity and collaboration. Different dimensions of proximity affect knowledge processes depending on the type of knowledge. A shared knowledge base is crucial for the process of collaboration in order to take advantage of proximity. Different forms of collaboration either correspond to exploration or exploitation of new knowledge. As a consequence, in the following master thesis we want to focus on the influence of proximity on collaboration.

The concept of proximity in relation to collaboration

The influence of different forms of proximity on the collaboration in nanotechnology networks was addressed in a few studies.

Most of the researchers are consistent concluding that geographical proximity positively affects research collaboration(Cunningham & Werker, 2012)(Katz & Martin, 1997). The success of different industry and academic clusters shows a support for this claim. Being spatially closer to each other increases the chances of getting acquainted and sharing ideas which can stimulate innovation. However, as it was shown above this form of proximity has to be complemented by other forms.

Findings of Cunningham and Werker suggest that collaborations in Europe are not randomly distributed(Cunningham & Werker, 2012). However, they are directly affected by geographical and technological proximities as well as indirectly by organizational proximity. Physical distance together with regional aggregation matter for nanotechnology researchers when they engage in collaboration. Mutual knowledge has an optimum level above which the collaborative productivity of parties decreases. It is shown that academic institutions mediate the influence of shared knowledge on the intensity of collaborative relationships. Additionally, on the example of one country authors showed that Dutch knowledge production is geographically and technologically concentrated(Werker & Cunningham, 2011). Dutch success in nanotechnology is the result of the concentration of activities in the industry.

A study of one Italian university shows how different proximity dimensions affect knowledge flows between nanotechnology network actors and knowledge gatekeepers. In particular, organizational proximity seems to play a role in small local networks, and geographical have a positive influence on the establishing of the relationships between organizations(Petruzzelli, 2008).Other scholars also empirically indicate the importance of geographical as well as other forms of proximity for the scientific collaboration and innovation (Maggioni & et.al., 2007)(Frenken, 2010).

Conclusion of the literature review

Thus far, we were able to demonstrate the complexity of the processes in nanotechnology networks. These networks can be characterized by the high degree of interdisciplinarity and institutional diversity. Due to a dispersed knowledge the sources of innovation are located in the networks rather than organizations(Smith-Doerr & et.al., 1996). Knowledge production and application is complex and is affected by the collaboration and proximity of network actors.The character of these relationships is different and requires constant investigation.

Research collaboration is a hot topic and is acknowledged to be a key mechanism for the knowledge production. Its study probably will never lose its relevance because of the evolving networks. With the development of nanotechnology new disciplines fuse into it shaping its technological trajectory. These disciplines bring together new scientists and organizations as well as change existing relationships.

We focus our research on the influence of proximity on collaboration in German nanotechnology. German research system is institutionally fragmented which represents an interesting example for the research. Understanding the character of these relationships would help to make policy decisions aiming to foster aninnovation in nanotechnology. This process is important because nanotechnology is pervasive and might have the potential to drastically change our lifes.

Conceptual model

In the following we look at the concepts of proximity and collaboration more carefully. Investigating the nature of these concepts we provide their definitions and formulate hypotheses to be tested in our research.

Collaboration

It is very important to understand what is research collaboration before starting to look at it in relation to other concepts such as proximity. The concept is central to the research networks, especially ones with a multi-disciplinary character. It has the power to facilitate knowledge production and to develop emerging technologies. However, when one thinks about the word collaboration it is rather vague to define. So what is research collaboration?

Katz and Martin provide a comprehensive study of the research collaboration. They provided different definitions and categorized its different levels. In general,research collaboration is defined as *"working together of researchers to achieve the common goal of producing new scientific knowledge"*(Katz & Martin, 1997). We use this definition for our analysis.

> *Research collaboration* is defined as working together of researchers to achieve the common goal of producing new scientific knowledge.

However, there is a problem of how closely researchers have to work together in order to be counted as 'collaborators'. A weak definition would include any researcher who contributed to the research, and a strong definition would include only researchers who provided input to all main research tasks. But in reality it is very hard to measure an exact contribution of every researcher to the final output of a joint project. Scholars have to come up with certain assumptions in order to resolve this problem.Research collaboration has very ill-defined borders which are a matter of social convention and open to negotiation(Katz & Martin, 1997).

This conclusion illustrates the scope of the possible research that can be done to study this concept. It also adds complexity to its exploration and exploitation. It is hard to understand what kind of relationships have to be built in order to gain the maximum output from the mutual work. But it is not only ill-defined borders that add vagueness to the concept of collaboration.

Different levels of collaboration can be distinguished.Collaboration levels have to be recognized although it is people who collaborate. It may occur between research groups within a department, between departments within an institution, and so on including industry sectors, geographical regions and countries(Katz & Martin, 1997). What is also important to understand is that collaboration may be within a particular level as well as between different levels.Prefixes *intra*- and *inter*- correspond to these names respectively. Following from that it can also be unambiguouslyhomogeneous including *intra* or *inter* and heterogeneous including both.

Given the multi-faced nature of collaboration we impose certain limitations in order to be able to study it. This procedure would allow us to operationalize the concept and associate activities in a real worldwith it. This master thesis focuses on the inter-organizational collaboration excluding possible interactions between researchers within an organization. The first limitation excludes other levels of collaboration leaving only organizational level. The second limitation excludes collaborations within a chosen level. The choice is supported by the expressed interest of the author in this type of collaboration.

Co-authorshipis one of the most common ways to measure collaboration. Katz and Martin indicate that it is not a perfect measure of such a multi-level and complex concept, however it gives certain advantages. In particular, the method is:

- invariant and verifiable;
- relatively inexpensive and practical;
- statistically more significant allowing for a large sample size(Katz & Martin, 1997).

Moreover, in our research we argue that there is a distinction between initiation and facilitationof collaboration activities. We propose that different dimensions of proximity may affect the establishment of collaboration and its development differently. It requires a certain effort to get to know another researcher and to conduct a joint project together. However, there are a considerable amount of reasons why researchers might not continue to work together after completion of the project. It could be a personal mismatch or too great distance between parties, or discontent of one party over another due to the results of a mutual work and number ofothers. The maintenance and development of the relationships requires completely different skills and is affected by other factors rather than the establishment of them. Therefore, we divide collaborative relationships in two stages. First stage is the establishment of collaboration or its initiation. Second stage is the development of collaboration or its facilitation.

Proximity dimensions

In the following we discuss proximity dimensions that were chosen for the purpose of our research. The concept of proximity itself is rather vague without specifying what dimension of it are taken into account.Definitions given to its dimensions in scientific literature are overlapping and sometimes contradictory(Knoben & Oerlemans, 2006). Therefore, the first task is to define proximity concepts in order to reduce the ambiguity. Another question that arises is which dimensions are the most relevant to inter-organizational collaboration? Knobenand Oerlemans attempt to answer these questions.

The authors were able to show which proximity dimensions are most relevant for inter-organizational collaboration (IOC). It is possible to identify what combination of dimensions capture the majority of proximity effects on collaboration. The process was done by decomposingproximity dimensions and comparing the underlying meaning included by researchers.Figure **2** shows the most relevant proximity dimensions for inter-organizational collaboration. Geographical, organizational and technological proximities grasp all definitions used by scholars in the literature(Knoben & Oerlemans, 2006). It is seen that organizational proximity is the most ambiguous concept as it may represent cognitive, institutional, cultural and social proximities as well as combine them.

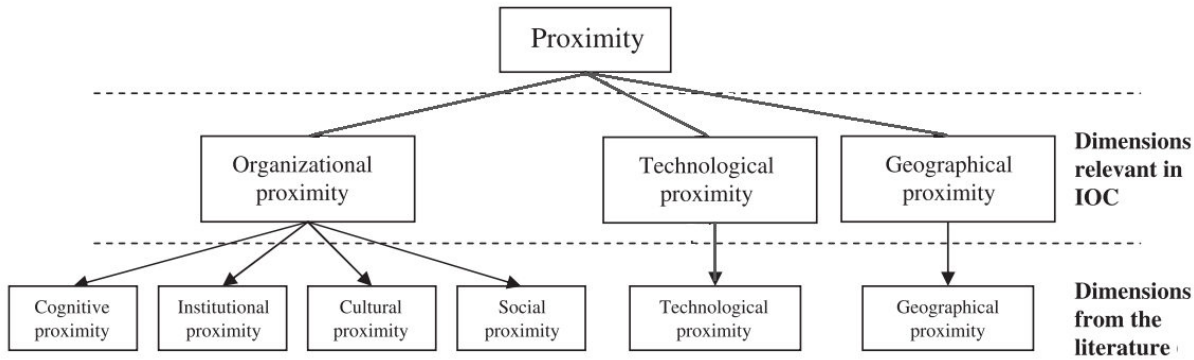


Figure 2 - Proximity dimensions relevant for inter-organizational collaboration (J. Knoben, 2006)

For the purpose of our research we choose three aforementioned dimensions because they account for the variety of different proximities. We add systemic proximity as a form of geographical proximity defined from the regional innovation systems points of view. All dimensions, their definitions and hypotheses formulated are presented below.

Physical proximity

Geographical proximity plays an important role in collaboration activities. This dimension is probably one of the most widely used in the scientific literature (Cunningham & Werker, 2012) (Werker & Cunningham, 2011) (Petruzzelli, 2008)(Knoben & Oerlemans, 2006).It is determined with the lowest level of ambiguity among scholars. The roots of the theories about the influence of geography on economic prosperity and innovation go back to the works of Alfred Marschall and his definitions of industrial districts(Belussi & Caldari, 2008).

Although, globalization and digitalization are seen as forces to eliminate the influence of geography on the research, there are arguments that it still matters (Morgan, 2004)(Porter M. , 1998). At first glance it looks like all the developments of ICT can substitute spatial proximity with virtual proximity. However, virtual communication imposes certain restrictions on the knowledge that can be transferred among partners. For example, a tacit knowledge is very much personally and context dependent being in a lot of times also locationallydependent. Often research requires thepersonal involvement of parties. For example, most of the time researchers have to be present in person while conducting different experiments. We say here that 'virtual' proximity to a greater extentcomplements physical proximity than replaces it. Spatial and virtual proximity are far from being mutually exclusive and co-evolve together (Morgan, 2004).

Face to face communication is very important in doing a mutual work.To know each other, researchers have to meet. Small distances between firms and organizations increase the chances that their employees might know each other. A 'silicon valley' might serve as a perfect example how spatial proximity may facilitate parties to engage in the collaboration which stimulates aninnovation and knowledge creation.

Some scholars emphasize the problem of mutual knowledge for the geographically dispersed teams of researchers (Cramton, 2001). Spatial proximity partly solves this problem. Facilitating collaboration helps parties to maintain the level of mutual knowledge needed for understanding each other. Otherwise if parties lack this communication due to a long distance between them the knowledge base may diversify to the extent that hampers a smooth transition of knowledge pieces.

Physical proximity is defined as geographical distance between two collaborators.

In this master thesis geographical proximity is called physical proximity because it represents a pure spatial distance between two collaborators.

We formulate two hypotheses relating the physical proximity. As it was mentioned earlier we are interested in the finding of initiators and facilitators of collaboration activities. Following from the discussion above we argue that both of these processes should be fuelled by physical proximity. Hence, first two hypotheses can be formulated as follows.

H1: *With the increase in physical proximity between two organizations the likelihood of them engaging in collaboration increases.*

H2: *The increase in physical proximity between two organizations facilitates their collaborative output.*

Systemic proximity

Systemic proximity is a share of common history, language, culture and communication patterns due to affiliation to a particular geographical region.

We added a systemic proximity in our research that takes into account the differences of geographical units. It is defined from the regional innovation system point of view. In any country regions have evolved following different trajectories shaped by political, economic, and other forces (Cooke & et.al., 1997). The characteristics of a regional research system are in addition to other reasons defined by a common history, language and culture. This can be reflected in a way universities, firms and other research institutions communicate and are organized. This ordering creates new routines, habits and norms and can also be called governance structure or 'social capital' (Cooke & et.al., 1997). Systemic proximity then takes place when organizations share this 'social capital'. In other words systemic proximity is a share of common history, language, culture and communication patterns due to the affiliation to a particular geographical region.

We argue that organization seeking for collaboration would more likely do that with the one that shares 'social capital' with it. This choice is motivated by the fact that such relationships would require less effort due to a common behavior and norms developed being in the same environment. We formulate two hypotheses regarding this dimension of proximity following our distinction between two stages of collaboration.

H3: *Systemic proximity between two organizations increases the likelihood of them engaging in collaboration.*

H4: *Systemic proximity between two organizations facilitates their collaborative output.*

Organizational proximity

Organizational proximity is the nextdimension considered. It suffers most from the high level of ambiguity in its definition(Knoben & Oerlemans, 2006).Exclusion of this dimension would not provide a full picture of the influence of proximity on collaboration. Here we first discuss possible definitions of organizational proximity and then explain which of them is used in this project.

First of all organizational proximity can be classified in two different levels: a structural level and a dyadic level (Knoben & Oerlemans, 2006). The structural level implies the equivalence in structural positioning of two organizations or their place in the same network with a particular properties. In other words it focuses on the extent to which the environments that surround organizations are overlapping. The dyadic level is focused on the structural differences or similarities inside organizations. It compares the context in which members of organizations operate.In the master thesis we define organizational proximity as an extent to which relationships are shared in an organizational arrangement (Boschma, 2005). Such definition includes both levels. Organizational arrangements can be shared by collaborators through affiliation to the same organizational structure as well as affiliation to organizations sharing the same environment.

Organizational proximity is an extent to which relationships are shared in an organizational arrangement.

Differences and similarities in organizational structures affect collaboration activities. Differences may impose constraintsfor researchers to collaborate. Similarities may facilitate the intention to collaborate with others.In particular, organizational structures of universities are expected to facilitate collaboration as they are seeking new connections. Researchers seek to apply or expand knowledge produced in university. The field of nanotechnology requires a synergy of different scientific disciplines and such open research infrastructures are expected to positively influence collaboration. Often the university system is organized in a way that scientists depend on the frequency with which their name appears in publications. A scholar may either publish more by himself or engage in joint research projects with others. The latter allows to perform more complex and therefore interesting research, solving the problem of publications. The German research system is characterized by a large share of 'non-university' research. Four main organizations, the Max-Planck Society (MPG), Helmholtz Research Centers (HGF), Fraunhofer Society (FhG) and Leibnitz Association (WGL), have developed quasi – functional monopolies in different research domains (Heinze & Kuhlmann, 2008). The research domains are fundamental research (MPG), applied contract research (FhG) or big-science research facility management (HGF). Researchers from extra-university sector may find it benefitial to collaborate with their peers from the same sector because they are working in similar organizations differentiating only in their research. Hence, we formulate two hypotheses aimed to explain the influence of organizational proximity on different stages of collaboration.

H5: *Affiliation of both organizations to the same university or 'non-university' sector increases their likelihood to engage in collaboration.*

H6: *Affiliation of both collaborators to the same university or 'non-university' sector facilitates their collaborative output.*

Technological proximity

A communication problems between collaborating parties may occur due to different knowledge bases. Technological proximity refers to the knowledge that two parties have. Nanotechnology involves two facts that require consideration in terms of knowledge. Firstly its sources are widely dispersed among network actors and secondly the knowledge itself is diverse due to the interdisciplinary character of the field.

Cognitive distance between employees shows the inverted U-shaped relationship with the innovation performance of a firm. This relationship is explained by Nooteboom. People see, evaluate and understand the world differently depending on the life paths and environment along which they developed (Nooteboom, 2007). Living their life they built a certain knowledge base. The degree to which a group of people shares such a base influences its productivity and capabilities. Sharing too much makes collaboration unproductive because none of the parties can complement another. Sharing too little imposes the problem of an understanding of each other.

Employees of a given organization have the same base determined by specifications of the organization. For example, in a mining firm one can find specialists in mining but not in tourist services. Moving now to inter-organizational collaboration we can expect the same tendency in peoples'affairs as in intra-organizational relationships.

Technological proximity is an extent to which organizations share knowledge bases and technological experience.

Therefore, technological proximityis based on the shared knowledge bases and technological experience. This definition makes use of the term relative 'absorptive capacity' which states that organizations should have enough similar knowledge base in order to understand each other but be different in specialized base in order to contribute to each othersexperience. In other words they have to establish and maintain a certain amount of the aforementioned mutual knowledge. A failure to do so can have serious consequences for the viability of collaboration (Cramton, 2001).

All of this indicates that a certain optimum level of shared knowledge exists. This level should be optimum for both stages of collaboration. If parties do not share enough knowledge base or have too overlapping knowledge base the likelihood that they will start a joint research project is low. The closer is shared base to the optimum level the higher is a chance of the establishing of the collaboration link. The same rule should apply when relationship between parties exists. Below a certain level technological proximity would facilitate collaboration activities. Conversely, above that level it starts to hamper collaboration because parties stop to understand each other. As a consequence we formulate the following hypotheses.

H7: *With the increasing technological proximity between two organizations their likelihood to engage in collaboration first increases and then decreases.*

H8: *With the increasing technological proximity between two organizationstheir collaborative output first increases and then decreases.*

Therefore we conclude that the influence of four dimensions of proximity on different stages of collaboration is analysed in this research project. This dimensions form a conceptual model of the project (Figure 3). The relationships between concepts are formulated in eight hypotheses that are tested in the following chapters.

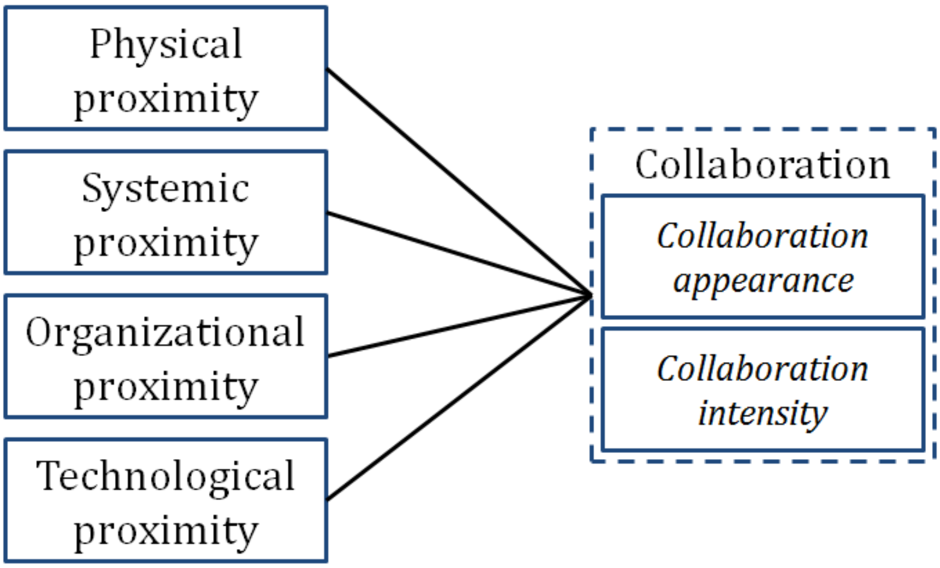


Figure 3 - The conceptual model

3. Research methodology

The following chapter presents the research methodology used for the analysis of German nanotechnology network. The methodology is aimed to test the hypotheses formulated in previous Chapter and to find evidence to support proposed arguments. It starts with explanation of research strategy, followed by the population and the sample of the study. Then it introduces the results of a sampling procedure. Finally, we explain the operationalisation process of four dimensions of proximity and present variables that are used in our research.

Research strategy

Due to nanotechnology peculiarities, namely interdisciplinarity and pervasiveness quantitative methodologies have to play a greater role in analyzing and forecasting this scientific field (Salerno & et.al., 2008). These techniques allow us to analyse bigger samples helping to cope with the evolution of extremely dispersed knowledge, complex relationships of actors and growing scientific networks.

The research strategy chosen is a quantitative, secondary research which will rely on existing data from publication databases.Using this approach we are able to createa picture of best German players in nanoscience field and their relationships.The bibliometric analysis of nanotechnology network is conducted measuring the concept of collaboration through co-authorship. This method provides certain advantages. It is invariant and verifiable, relatively inexpensive and practical, and statistically more significant allowing for large sample size (Katz & Martin, 1997).

Investigation of nanoscience publications provides us with the analysis of the latest theoretical developments available to all researchers. We can say that publications reflect the production of theoretical knowledge while patents reflect more practice oriented knowledge. Publishing is also one of the major ways of exchanging the knowledge. Of course not all knowledge produced is published.However, if knowledge is not documented it rarely contributes to the process of its sharing. And therefore, it does not add value to the development of the science.

In order to test developed hypotheses three regression models are used. The first model is looking at the influence of proximity dimensions on the likelihood of collaboration appearance between two actors. This model is specifically aimed at testing hypotheses number 1,3,5, and 7. The second model is predicting the intensity of collaboration assuming that a link between actors already exist. It is aimed to test hypotheses number 2,4,6 and 8. The third model is predicting both the appearance and the intensity of collaboration. The model is aimed to allow the comparison with the work about European nanotechnology (Cunningham & Werker, 2012) and test the results of the first two models.

Data collection methods

The data were downloadedfrom Web of Knowledge database of Thomson Reuters (Thomson Reuters, 2012) using a specially designed query. The choice of the query as well as the database is explained below.

While conducting a bibliometric analysis of an emerging technology it is necessary to apply the definition of this technology.It allows to exclude from the study publications containing only non-relevant words like “$NaNO_2$” and others. Over the past years there were several studies aimed to obtain a bibliometric definition of the nanotechnology (Porter & et.al., 2008)(Dang & et.al., 2012)(Grieneisen & Zhang, 2011)(Kahane & Mogoutov, 2007). The work is tangled by the fact that technology itself develops and includes new fields as well as new words like ‘graphene’. We use the work of Arora, Porter, Youtie, and Shapira published in 2012 (Arora & et.al., 2013). This query is an update version of a search strategy developed in 2006. And hence,it captures the latest achievements in nanoscience and technology.

The approach of this so-called evolutionary lexical query comprises a modular key word search strategy with a two-step inclusion and exclusion process(Arora & et.al., 2013)(Huang & et.al., 2011). A semi-automated search process was applied to discover the trending keywords which were accessed by the industry experts from different fields. The query was already used in other research papersin its original version and showed its reliability.

Web of Knowledge database was chosen as a source of a publication data. The rationale behind it is that it covers world’s most important and influential journals including 12000 of top tier international and regional journals (Thomson Reuters, 2012). The data provide information about abstract, publisher, authors and their addresses, citations, funding, research place and etc. An example of a publication information extracted from the database can be seen in Appendix A. An additional factor in favor of this database is that the aforementioned query was tested and proved its reliability on the data from the Web of Knowledge.

The time span of publications is from the year 2010 to 2012 inclusive. Because the query was focused on the latest progress in nanotechnology such a choice captures all latest published achievements. Additionally, a bigger time span will add ambiquity to the research because German research organizations change over time by dividing and merging. For example,Karlsruhe Institute of Technology (KIT) was founded by a merger of Research Center Karlsruhe and University of Karlsruhe in 2009.

The analysed database contained more than 270000 publications extracted from Web of Knowledge with the help of the chosen query. Around 20000 of them had at least one author affiliated to German organization. These 20000 publications served as a basis for building a network of inter-organizational collaboration.

Population of the study

All German organizations working in the field of nanotechnology are the research population of this study. The geographical limits of the population are clear; it has to be located in Federal Republic of Germany. More precisely it has to have at least one of its research centers or departments involved in nanotechnology research and located in Germany.

In the Web of Knowledge database authors are affiliated to organizations. Most of the organizations are providedwith their addresses excluding rare exceptions. These address served as a basis for imposing the geographical limits on the population identifying organizations located in Germany.

Here we assume that the work published was performed in the places to which its authors are affiliated.

The second limit of the population is that the organization has to conduct a research in nanotechnology. Applying NST definition tells us that organization should have a manipulation of a matter having at least one dimension from 1 to 100 nanometers. Although such definition probably fits to a lot of organizations not all of them are actively involved in the knowledge production and exchange processes. Relying on a chosen query we assume that all organizations that have their employees among the authors of any extracted publication are performing the research in nanotechnology.

Sample of the study

The most important German organizations are identified as a sample for the research project.A chosen sampling strategy is opting for the depth of the research and therefore, doesn’t include all German organizations. Such decision is also supported by the fact that including all organizations in the research might not be feasible. It may first lead to the information overload during the analysis. Secondly, cutting out the most productive organizations allows the reduction of the level of noise in the data that would inevitably occur otherwise. And thirdly, it guarantees that population assumption made because of the choice of bibliometric analysis will not be violated. Here we mean that it is feasible to check for 100 organizations that their research facilities are located in Germany and they do perform research in nanotechnology.

Two steps of the sampling procedure can be distinguished. First step is implicitly embedded in the choice of the bibliometric analysis. Only the organizations whose employees publish their work were included in the study. The advantages and disadvantages of this method were discussed above. It should be said that the contribution of organizations that do not publish their findings can be considered negligibleto the whole process of knowledge production and exchange through collaboration.Therefore it is not relevant for the current study.

The second step is the identification of the most productive organizations in terms of their publication output.Sampling by organizations provides an opportunity to account for failed collaborations in the research. This method allows us to see the influence of proximity on the establishment of relationships between partners. The criteria for identifying top organizations was the publication productivity which is a solid representation of its research activity. First, a contribution of each author to the paper was distributed equally. Second authors were allocated according to their organizations. A negligible amount of authors had two or more affiliations. Then each organization was ranked according to the number of publications produced by its authors. Therefore, if a paper is written by three authors from two organizations one will get one third and the other will get two thirds of a ‘weight’. Figure **4** shows an example of how the publication’s ‘weight’ was distributed among collaborators.

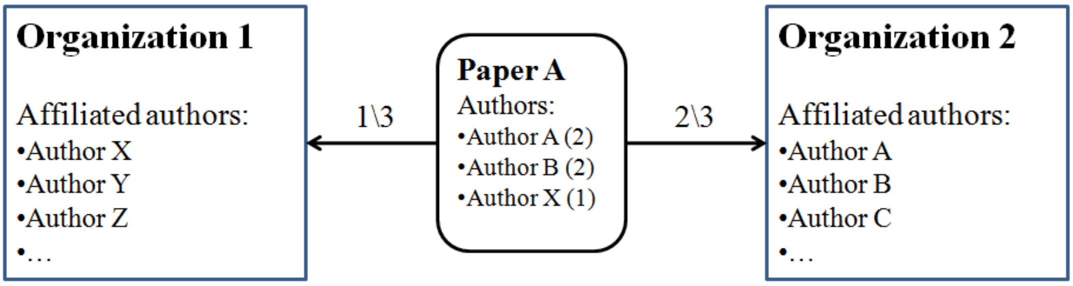


Figure 4 – Calculation of publication productivity of organizations

Results of the sampling procedure

As a result of the sampling procedure 100 most active in NST German organizations were identified and were included in the constructed network. In total more than 2000 German organizations were found that have at least one affiliated author in at least one publication. The top 100 organizations take part in 90 per cent of all publications by German organizations and in around 24 per cent of more than 270000 publications around the world. These organizations were chosen as a sample for a further research as the most productive according to the criteria described above.A complete list of the first 100 organizations can be found inAppendix C.It has to be said that during the calculation of their productivity the reorganization of the data was made accounting for the different spelling of organization's names. The work was done manually with the help of a software written in Python programming language. Some universities had more than 15 different ways of spelling; for example, Technical University of Munich had ten variations of its name. An example is shown in Table 1. Some examples of multiple spelling are presented inAppendix B.

Table 1 - Spelling variations of a university name in theWeb of Knowledge database

Number	Spelling
1	Tech Univ Munich
2	Tech UnivMunhen
3	TechnolUniv Munich
4	LS AC Tech UnivMunchen
5	Tech UnivMunchenPhys
6	TUM
7	Tech UnivMuenchen
8	TU Muenchen TUM
9	Tech UnivMunchen TUM
10	TU MunchenInstAdv Study

Figure **5** shows the share of organizational types in the sample. The majority of them are universities. Extra-university sector accounts for 38 per cent of all organizations. There is only one firm with non-academic background: BASF SE. It has the 69th position in our ranking with a marginal contribution to German publications equal to 0.43 per cent.

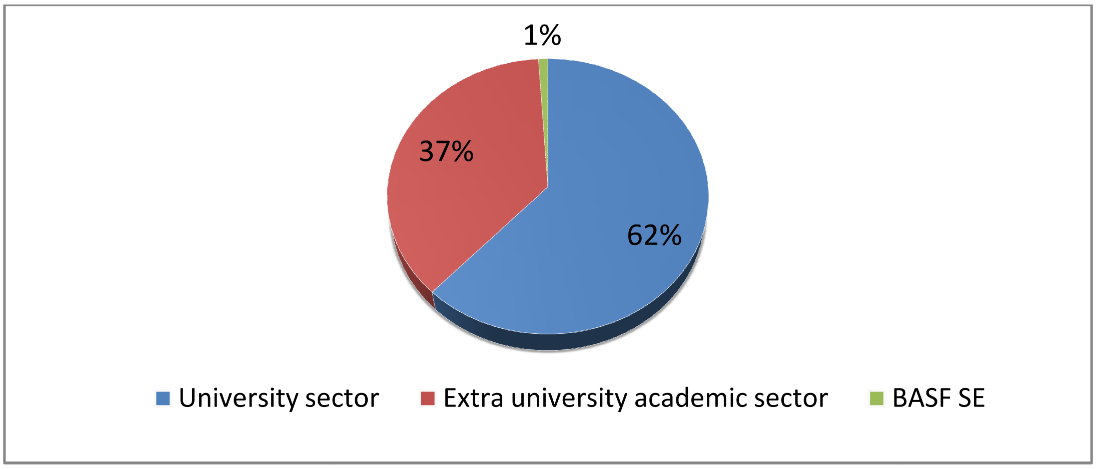


Figure 5 - German Nanotechnology Organizations

The Top 10 organizations account for a quarter of all publication output in nanotechnology. We can see here the importance of the extra-university sector in the German research system. Even among top 10 organizations there are three research centers as shown in the Table 2. It is notable that there is a representative of every major research association except Fraunhofer Society (FhG). First Fraunhofer Institute for Applied Solid State Physics appears only at 84th place in our ranking. FhG institutes mostly conduct contract research for firms and public agencies. It can explain their low publication output. Their research is mostly reflected in patent publications and no published R&D.

We spot some dynamics in the leading positions in German nanotechnology. According to Heinze in 2001, University of Hamburg, Max Planck Institute for Metal Research, Technical and Ludwig Maximillian Universities of Munich together with Research Center Julich were among the top European organizations in terms of yearly publication output in nanotechnology (Heinze, 2004). Collected data show that three of them except University of Hamburg and Max Plank Institution stayed in the top 10 most productive organizations in Germany. The reason for such adifference can layin the dynamicsof nanotechnology network and the fact that new organizations are taking lead in research. Another reason might lie in the differenceof queries used for downloading publications.

Table 2 - Top 10 German nanotechnology organizations

Number	Name of an organization	Ranking coefficient	Cumulative percentage of publications covered
1	University of Karlsruhe TH	676.4	3.67%
2	University of Erlangen Nurnberg	573.5	6.70%
3	Technical University of Munich	448.2	9.67%
4	Ludwig Maximillian University of Munich	416.6	12.62%
5	Helmholtz ZentrumJulich	395.0	15.32%
6	Technical University of Dresden	381.3	17.78%
7	RWTH University of Aachen	317.3	20.05%
8	University of Essen Duisburg	316.8	22.03%
9	Max Planck Institute for Polymer Research	310.2	23.85%
10	Leibnitz Institute for Solid State & Materials Research	299.2	25.87%

The network of aforementioned 100 organizations served as a sample for the study of the influence of proximity dimensions on collaboration. As it was mentioned above two organizations were considered as collaborators if they have at least one publication where authors affiliated to them are listed together. The network is presented on Figure **6**. The colour and the size of each node indicates its importance in this network. The importance was calculated using the eigenvector centrality measure from social network analysis theory. For more information see the following reference (Ruhnau, 2000). The thickness of the line between organizations depends on the amount of mutual publications. This amount was calculated accounting for the share of authors from two collaborators among the total number of authors of a particular publication.This fractional publication approach minimizes an error in estimating the input of every collaborator to the paper. However, a small error is still present based on the assumption that everyone contributes equally to the work done. Unfortunately, it is almost impossible to account for this bias.

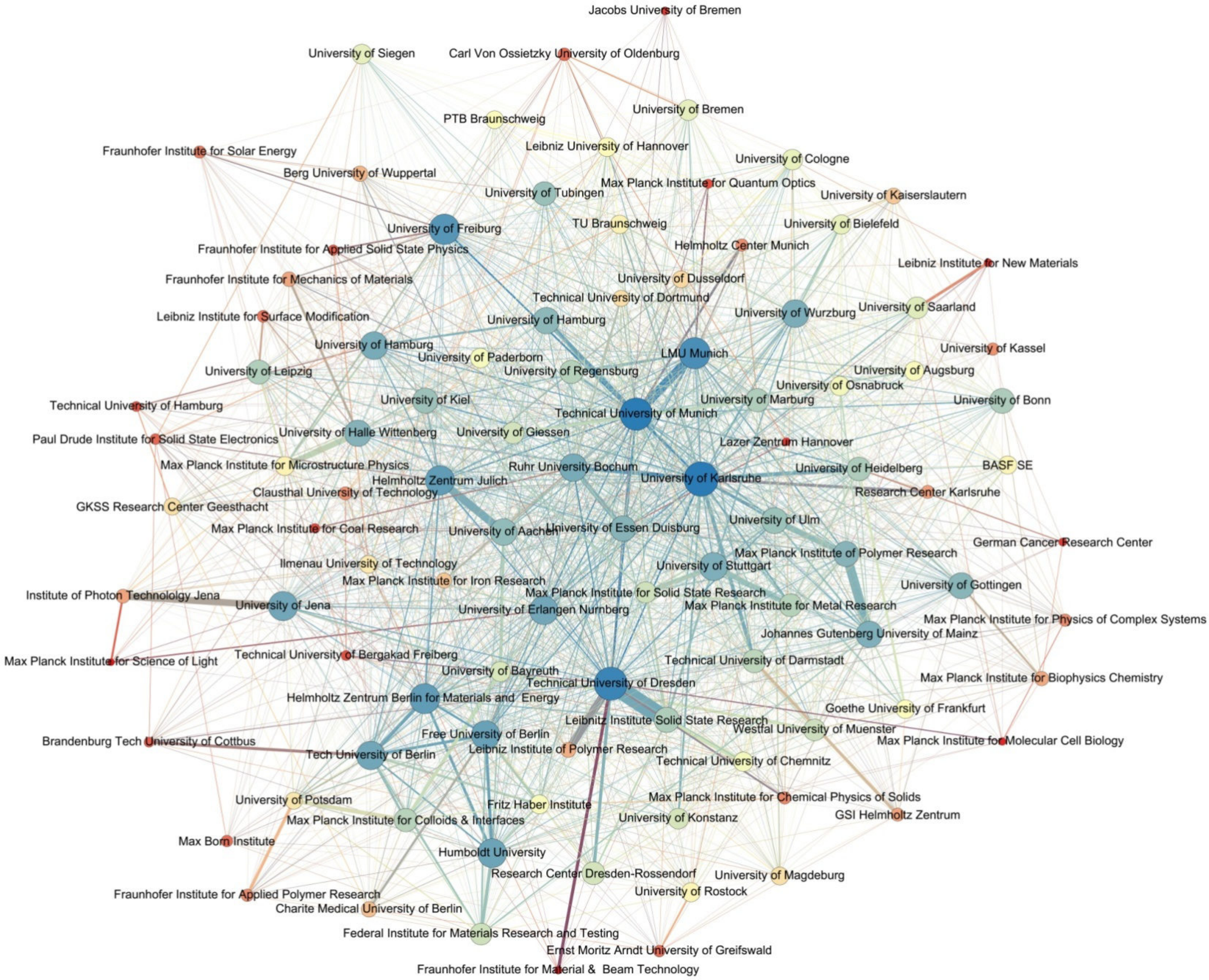


Figure 6 – Constructed Network of 100 German Organizations

We find two resons for the poor representation of firms in our sample. The first reason is the lack of resources to conduct a nanotechnology research. An organization needs to employ a large amount resourses in order to be actively involved in the network of organizations that represent nanotechnology research frontier. Second reasonfollows from the choice of the publication analysis. Publications mainly reflect knowledge production processes while firms are usually market oriented and focus on knowledge application processes which are better reflected in patents. Finally various barriers for publishing the research of the firm can be a third reason. A scholars may not be allowed to disclose their findings due to confidentiality issues or lack of time for publication preparation. Because of theses reasons our network presents an academic collaboration between organizations omitting non academic collaboration.

The hypotheses formulated in Chapter 2 were tested using the constructed network. The following section introduces all independent variables used in three constructed regression models. All variables were the same for the each model. Only the dependent variable and the model itself were different.

Research variables

The following part explains the operationalisation process and introduces independent variables included in the regression models.Four dimensions of proximity from the conceptual model developed in Chapter 2 were operationalized into eleven variables excluding interaction terms between them. Two types of variables were used: relational and non-relational. Relational variables are dependent on collaboration activities of two organizations while non-relational variables are independent from them and represent properties of a particular organization.Dependent variable in every model represents the concept of collaboration. The operationalisation of it is done differently depending on the objective of the regression model and is not discussed in this chapter.

Physical proximity variable

Physical proximity is represented by the logarithm of the distance between collaborating organizations. The geographical information was obtained using thelatitude and longitude data provided by Google Maps service. Here we took the addresses of main offices from the websites of organizations. The formula used for the calculation of the distance is shown below:

$$\mathrm{D} = \arccos[\sin(\frac{Lat1}{180^0}\bullet\pi)\bullet\sin(\frac{Lat2}{180^0}\bullet\pi) + \cos(\frac{Lat1}{180^0}\bullet\pi)\bullet\cos(\frac{Lat2}{180^0}\bullet\pi)\bullet\cos(\frac{Lon2}{180^0}\bullet\pi - \frac{Lon2}{180^0}\bullet\pi)]\bullet R$$

Where D is distance, *Lat* and *Lon* are latitude and longitude of a first and second organization respectively and R is the radius of the Earth equal to 6378.137 kilometers.

Systemic proximity variables

Systemic proximity is represented by six dummy variables indicating a shared or bordered NUTS region of Germany. The abbreviation means *Nomenclature of territorial units for statistics*.Thisclassification is a hierarchical system for dividing up the economic territory of the European Union. All affiliations were obtained through the website of the European Commission and the classification is valid until December 31st, 2014 (The European Commission, 2012).

There are two reasons for this administrative division. Firstly, in Germany federal states or *Landers*(German)are to a great extent independent in their policy decisions especially on the NUTS1 level. This allows us to expect a benefit for two collaborators being in one system or shared region. And secondly, the European nomenclature reflects cultural and historical background what goes hand in hand with a given definition of systemic proximity.

Organizational proximityvariables

Indicating the difference in organizational structure between universities and 'non-universities' organizational proximity was represented by two dummy variables. These variables in combination are aimed to show if a certain organizational structure facilitates collaboration. One indicates whether both of the collaborators are universities (*University*), and other indicates whether both of them are 'non-universities' (*Non-University*). It allows to account for all possible combinations of collaborators. Table 3 shows the coding of the variable.

Table 3 - Dummy coding of organizational variables

Type of collaboration	Abbreviation	Variable	
		University	'Non-University'
University - University	Univ - Univ	1	0
Non-university – Non-university	NUniv - NUniv	0	1
Mixed collaboration	-	0	0

Technological proximity variable

A research profile of every organization was collected in order to define the technological proximity or the level of shared knowledge and research experience. This research profile bases on the Science Categories of Web of Knowledge. In the following we explain how this approach suits to the provided definition of technological proximity.

These categories indicate scientific disciplines to which provided in a publication knowledge belongs. We see these categories as a discrete measure of a knowledge base. Each organization can specialize more in one of them and less in the other. We imply that if authors from that organization publish most of their papers in particular science categories it can be an indication of organization's specialization. Therefore we indirectly identify the knowledge base. In order to spot a shared knowledge base we focus on joint publications again looking at the likelihood of them to publish in a particular science category.

A possible weakness of this approach is the assumption that Web of Knowledge is allocating the papers according their real meaning. Providing the reputation of Web of Knowledge and that thousands of employees and peer-reviewers are working there this threat can be considered negligible.

Six main categories for nanotechnology were identified according to the query chosen. Table 4 shows the amount of publications per each category in the database analysed.

Table 4 - Number of publications in six major nanotechnology science categories

Science Categories	Number of publications
Materials Science, Multidisciplinary	83660
Chemistry, Physical	56279
Physics, Applied	53906
Chemistry, Multidisciplinary	49344
Nanoscience& Nanotechnology	45839
Physics, Condensed Matter	36464

We calculated the share of publications of an organization falling into a particular category. The research profile of each organization included the amount of publications in each of these categories plus 'all other' category. In addition to these profiles we calculated a profile for each of the 4950 possible collaborations between 100 organizations. This profile showed the amount of joined publications falling in one of the seven categories. In theory, the technological distance can be equal to infinity if organizations do not share publications in any of the categories. Here the inclusion of 'all other' category ensures a minimum level of technological proximity.

The same approach to the measure of a mutual knowledge can be found in the work of Werker and Cunningham (Cunningham & Werker, 2012). The variable is called mutual information and is calculated as follows:

$$X = \sum_{i \in X} \sum_{j \in Y} p(x_i, y_j) \log\left(\frac{p(x_i, y_j)}{p(x_i) p(y_j)}\right)$$

The formula is known from the theory of mathematical communication (Shannon, 1948) and involves the profile of the research done individually by the organization (i.e. *p(x), p(y)*) as well as the profile done mutually by two organizations (i.e. *p(x, y)*). Mutual information doesn't specify the basement of the logarithm. Here,a used base of the logarithm is ten.

One non-relational control variable is additionally included in regression models, namely total number of publications of a pair of organizations. The control variable helps to check the assumption whether collaborations are distributed according to the productive output of organizations.

Interactions between research variables

Apart from the variables described above a number of interaction terms are included in the models. First is a square of mutual information that represents technological proximity. The interaction is included in order to control for the possible non-linear effect expected according to the literature review. Secondly, universities and 'non-universities' are expected to mediate technological proximity. The role of mutual knowledge in university collaborations may be less than in extra-university collaborations or vice versa. Hence, interaction terms between organizational and technological variable are introduced.

Table 5 presents all variables and their descriptions. A complete list of organizations with all complemented information can be found inAppendix C.

Table 5-Independent regression variables and their interactions

Variable	Name	Description
Geographical variables		
X_1	Distance	Logarithm of distance between collaborators
X_2	Shared NUTS1	Shared NUTS1 region
X_3	Shared NUTS2	Shared NUTS2 region
X_4	Shared NUTS3	Shared NUTS3 region
X_5	Bordered1	Bordering NUTS1 regions
X_6	Bordered2	Bordering NUTS2 regions
X_7	Bordered3	Bordering NUTS3 regions
Technological variable		
X_8	Mutual Information	Mutual information about research profiles
Organizational variables		
X_9	Univ - Univ	Both of the collaborators are university
X_{10}	NUniv - NUniv	Both of the collaborators are 'non-university'
Control variable		
X_{11}	Publication	The average of the logarithm of the total number of publications
Interaction terms		
X_{12}	Square of Mutual Information	Square of Mutual Information about research profiles
X_{13}	Mutual Info * Univ	Interaction of university and mutual information
X_{14}	Mutual Info * NUniv	Interaction of 'non-university' and mutual information

In addition to the constructed network of collaborations, the data of 100 organizations were complemented with the information about their geographical location and affiliation of every organization to three NUTS regions: NUTS1, NUTS2, and NUTS3. In addition every organization received an indication whether it is a university or 'non-university'. And finally research profile of each organization included the amount of publications in every chosen science category plus 'all other' category.

Descriptive statistics of research variables

In the following section we present the descriptive statistics of all regression variables together with their collinearity.

The descriptive statistics shows that there are 58.1 per cent of failed collaborations in the analysed network. Maximum collaboration intensity can be spotted between Technical University of Dresden and Leibnitz Institute for Solid State Research and its value is 102. 83.3 per cent of collaborations including failed have the intensity below the value of 1 and 99 per cent of them have it below the value of 11. This fact indicates a high degree of co-authorship in nanotechnology and the involvement of third parties in the mutual work of organizations.

The geographical data suggest that the average distance between German collaborators is around 310 kilometers. Looking at thesystemic proximity 8.3 per cent of partners are located in the same Federal State (*Bundesland*), 3.2 per cent share a NUTS2 region and even less 2.1 per cent share a NUTS3 region. Although almost 26 per cent of collaborators are situated in the bordering regions on the first NUTS level, very few of them have bordering NUTS2 and NUTS3 regions: 9.6 per cent and 1 per cent respectively. The results are overlapping with the results of Werker and Cunningham were the majority of inter-organizational collaborations in Europe occurred in bordering NUTS1 and NUTS2 levels(Cunningham & Werker, 2012).Table 6 presents results described above.

Table 6 - Systemic proximity in German nanotechnology network

	Shared NUTS1	Shared NUTS2	Shared NUTS3	Bordered NUTS1	Bordered NUTS2	Bordered NUTS3
Percentage of occurred collaborations in the sample	8.3%	3.2%	2.1%	26%	9.6%	1%

The results of the composition of the network are consistent with old findings of Heinze. In 2008 a striking feature of German research system was a large share of extra-university public sector research with the majority of collaborations occurring between universities themselves and the public sector.While very few collaborations occurred within the extra-university sector (Heinze & Kuhlmann, 2008).We spot the same tendency in 2010 – 2012. 38.2 per cent of inter-organizational activities occur between universities and 47.6 per cent between universities and 'non-universities', while the remaining 14.2 per cent include collaborations within the'non-university' sector.

Appendix D presents more comprehensive descriptive statistics of all categorical and continuous variables used in regression models.

Multicollinearity

The variance inflation factor was calculated for each of the independent variables and their interactions in order to check for multicollinearity between them. This technique indicates how much the variance of an estimated coefficient is inflated due to thecollinearity. Constructed variables as well as their interactions did not show a significant variance inflation. All factors appeared to be below the value of 5 which can be considered as negligible(O'Brien, 2007). It excludes the possibility of inflated regression coefficients and allows to rely on their values.Table 7 presents all variance inflation factors.

Table 7 - Variance Inflation Factors of research variables

Variables	Collinearity Statistics	
	Tolerance	**VIF**
Distance	.230	4.342
SharedNUTS1	.333	3.000
SharedNUTS2	.206	4.865
SharedNUTS3	.265	3.770
BorderedNUTS1	.709	1.411
BorderedNUTS2	.482	2.073
BorderedNUTS3	.799	1.251
MutualInformation	.558	1.791
Square of MutualInformation	.399	2.507
Univ - Univ	.424	2.356
NUniv - NUniv	.468	2.137
Publication	.864	1.157
Mutual Information*Univ	.319	3.135
Mutual Information *NUniv	.448	2.234

4. The influence of proximity on the likelihood of collaboration

The following chapter presents the first regression model. First, it explains the dependent variable and presents the equation of the model. Afterwards it shows the fulfillment of its requirements. Finally, the results of the regression are analysed. This regression model is developed to test hypotheses about the influence of proximity dimensions on the establishment of collaboration.A link formation was investigated for firms showing that accumulated knowledge base increases the chances to find a partner for joint research (Ahuja, 2000). In our project in addition to knowledge base we also look at the influence of shared 'social capital', physical proximity and organizational structure.

Dependent variable

The concept of collaboration is represented by the dependent variable. The variable is binary indicating the existence and the abcense of the link between two organizations. If the number of publications between two organizations is zero the variable is also zero indicating a failed collaboration. If there is at least one publication with authors affiliated to both organizations the variable scores 1, indicating successful collaboration. All 4950 possible links between the top 100 organizations are taken into account in the model.

Model equation

For the purposes of the first modela binary logistic regression was chosen which specification can be expressed as follows:

$$Y(X_1, ..., X_n) = \frac{e^{(\alpha + \beta_1 X_1 + ... + \beta_n X_n)}}{e^{(\alpha + \beta_1 X_1 + ... + \beta_n X_n)} + 1}$$

Where Y is a binary dependent variable, α is an intercept, β is a coefficient of independent variable and X_i are independent variables.

The regression is used to model the influence of different factors on a dichotomous variable.It shows the likelihood of the dependent variable falling in one of the categories;in this application it shows the likelihood of organizations to engage in collaboration and aims to identifyproximity dimensions that are crucial for the establishment of a link between the researchers.

One of the indisputable advantages of the logistic regression is that it does not require typical assumptions of normality, linearity and/or homoscedasticity. So for example, the estimated error is not necessarily normally distributed. However, the model requires a linear relationship between the logit of independent variables and dependent variable together with relatively 'large' sample size. The sample of 4950 cases can be considered sufficient.

We used a two step logistic regression in the research. It means that the model including all independent variables is compared to the intercept only model. In the first step only a constant is included. In the second step all independent variables are added to the constant after which the significance of the difference between two steps is calculated.

Specifications of the application of thefirst model

The binary logistic regression was significant with the 99 % confidence interval.The dependent variable is dichotomous. The first step of the model included only an intercept. On the second step the model included all independent variables as well as their interactions. Table 8 shows the increase in the model prediction capacity from 58 % to almost 91%. Here we report the Cox & Snell pseudo R square of the model which is equal to 0.576. Its value shows the improvement of the full model over the intercept only model. The Nagelkerke R Square test shows that our model explains 78 % of the variation in the data.

Table 8 - The prediction improvement of the logistic regression model

	Observed		Predicted		
			Dependent		Percentage Correct
			0.00	1.00	
Step 1[a,b]	Dependent	0.00	2878	0	100.0
		1.00	2072	0	0.0
	Overall Percentage				**58.1**
Step 2	Observed		Predicted		
			Dependent		Percentage Correct
			0.00	1.00	
	Dependent	0.00	2649	229	92.0
		1.00	226	1846	89.1
	Overall Percentage				**90.8**

a. Constant is included in the model.
b. The cut value is .500

Regression results of the first model

We first remind which hypotheses are tested using the model deleveloped before the start of the analysis of the results of the model execution. The binary logistic regression tests four out of the eight formulated hypotheses. All of them are presented below.

H1: *With the increase in physical proximity between two organizations the likelihood of them engaging in collaboration increases.*

H3: *Systemic proximity between two organizations increases the likelihood of them engaging in collaboration.*

H5: *Affiliation of both organizations to the same university or ' sector increases their likelihood to engage in collaboration.*

H7: *With the increasing technological proximity between two organizations their likelihood to engage in collaboration first increases and then decreases.*

Table 9 presents the results of coefficient estimation for the binary logistic regression. Only the variables representing systemic proximity showed insignificant coefficients. All others including intercept appeared to be significant.

Table 9 - Binary logistics regression coefficients

Independent variables	B	S.E.	Wald	df	Sig.	Exp(B) Odds ratio
Distance	-.231	.120	3.694	1	.000	.794
Shared NUTS1	.341	.333	1.050	1	.306	1.407
Shared NUTS2	.853	.653	1.706	1	.192	2.347
Shared NUTS3	-.176	.709	.062	1	.804	.838
Bordered NUTS1	.282	.157	3.226	1	.072	1.326
Bordered NUTS2	.091	.279	.106	1	.745	1.095
Bordered NUTS3	-.342	.568	.362	1	.547	.710
Mutual Information	2.777	.234	140.539	1	.000	16.076
Univ - Univ	3.172	.375	71.580	1	.000	23.857
NUniv - NUniv	-1.120	.240	21.778	1	.000	.326
Square of Mutual Information	-10.197	.471	467.792	1	.000	.000
Mutual Info *Univ	-6.549	.746	77.001	1	.000	.001
Mutual Info *NUniv	2.258	.583	15.029	1	.000	9.567
Publication	2.369	.131	325.435	1	.000	10.685
Constant	2.608	.147	314.404	1	.000	13.569

The results suggest a support for the hypothesis 1. In particular, the reduction of physical distance between organizations facilitates the establishment of collaborative relationships. The closer are organizations the more likely they will engage in collaboration as can be seen from a significant negative coefficient of a distance in the first model. The odds ratio in the last column tells us that an increase in the distance reduces the likelihood of collaboration by around 20 per cent.The result is in accordance with conclusions of other scholars confirming widely accepted opinion that geographical distance still plays an important role in the collaboration despite technological developments.

No support was found for the hypothesis 3. Due to the fact that German states have their own research and funding policy a segmentation of the nanotechnology network was expected. However, regionsdo not show a significant influence on collaboration processes. From our model we can see that collaborative links between organizations are distributed regardless affiliation to NUTS regions. The reson for that can be various German nanotechnology initiatives and coordination of activities(Federal Ministry of Education and Research, 2012) that successfully eliminated the role of regional differences in the joint research of organizations.

A contradictory results were found for the hypothesis 5. On the one hand the likelihood of collaboration between universities prevails on mixed collaborations. In particular, a presence of two universities in a pair of organizations boosts the odds of establishing a link between them. On the other hand extra-university sectors shows negative influence on collaboration. A presence of two ‘non-universities’ in a pair of collaborators decreases the chances of successful relationships. Therefore we can not accept hypothesis 5.

We find several reasons for our results.Collaboration between universities allows them to share costs of their research(Lavie & Drori, 2012)enchance its visibility or expand and improve the equipment and facilities (Heinze & Kuhlmann, 2008).Institutes from a large extra-university research sector in Germany haven’t traditionally collaborated with each other and were organized according to specializations (Heinze & Kuhlmann, 2008). For example, Max Planck Society traditionally focused on free basic research in innovative fields while Fraunhofer Society concentrates its research efforts on application oriented research(Federal Ministry of Education and Research, 2012). We emphasize here that organizations do not specialize according to scientific disciplines but specialize according to the purpose of their research. Such specialization can explain a negative effect of 'non-university' collaboration. Here we can refer to a famous Pasteur's Quadrant separating basic, user-inspired and pure applied research. One of the main differences between these research types is their objective. Institutes are less eager to engage in collaboration if they have differing objectives.We see that nanotechnology network in Germany is not only divergent in terms of disciplines and institutions but also there is a diversity in objectives of these institutions.

Another possible explanation we find in the work of Heinze(Heinze & Kuhlmann, 2008).Authors find several barriers to inter-institutional research collaboration through interviews. First is stereotypes and prejudices when researchers from some organizations are considered to be slower or less productive. Second is incompatible working routines rooted in different missions of various institutes. And third is the lack of an interface management among headquarters of organizations. These reasons could influence the likelihood of establishing a link between parties.

Additionally, organizational proximity mediates the effect of technological proximity on collaboration initiation. It can be seen through interaction term among organizational and technological variables. We show that when 'non-universities' initiate collaboration the role of mutual information increases, while in case of university collaboration it decreases. Due to the aforementioned misalighment of objectives in extra-university sector a shared knowledge base becomes a reson to collaborate for 'non-university' organizations. Contradictory, in universities researchers probably try to find new combinations of knowledge and are more eager to engage in different collaborations.

We accept hypothesis 7. The positive effect of technological proximity on the initiation of collaboration can be seen in the positive coefficient for the variable mutual information. The negative effect is reflected in the coefficient for the square of this variable. Therefore, we show that shared knowledge base has an inverted U-shaped relationships with the chances of establishing a successful collaboration between two nanotechnology organizations. Our result is in accordance with the theory about optimal cognitive distance or shared mutual knowledge (Nooteboom, 2007) (Boschma, 2005). In our network organizations engage in joint research only if they publish in similar science categories which means that they conduct similar but not the same research.

Additionally, we indicate that an increase in publication productivity of two organizations boosts the chances of them to collaborate. An increase in publishing rate raises the awareness of the research being done and attracts attention of other scholars. We can see here that nanotechnology research is mainly conducted in collaboration with other parties. If it would be done mainly internally the odds ratio for this variable would be much smaller.

Finally, we conclude that physical distance, presence of a university, mutual knowledge base and publication productivity are factors that increase the chances of two organizations to engage in collaborative nanotechnology research. We showed that systemic proximity plays no role in a choice of a partner for a German organization. Extra-university sector continues to hamper collaboration among its parties. After looking at the likelihood of collaboration we move to the investigation of how the same factors affect the intensity of established links.

5. The influence of proximity on the intensity of collaboration

The following chapter presents the second regression model. After looking at the likelihood of engaging in collaboration the second regression model is aimed to test hypotheses regarding the influence of proximity dimensions on the intensity of already established relationships between organizations. First, the chapter explains the dependent variable and presents the equation of the model. Afterwards it shows the fulfillment of its requirements. Finally, the results of the regression are analysed.

Dependent variable

The concept of collaboration is represented by the dependent variable. The variable was calculated using the fractional publication and is equal to the total amount of publications between two collaborators.In the model the link between researchers is assumed to exist. If the number of publications between two organizations is zero this collaboration is not taken into account in the following model.The variable is ratio scaled. 1994 collaborations were found in the constructed German nanotechnology network.

Model equation

Amultiple linear regression was selectedand its equationis specified as follows.

$$Y(X_1,\dots,X_n) = \alpha + \beta_1 X_1 + \dots + \beta_n X_n$$

Where Y is adependent variable, α is an intercept, β is a coefficient of independent variable and X_i are independent variables.

However, if logistic regression is tolerant to the assumptions of normality, linear regression requires their fulfillment. The estimated error is expected to have normal distribution. Dependent variable has to have approximate linear relationship with independent variables both individually and grouped. Importantly, no independent variable can be a linear combination of other independent variables. This is called the issue of multi-collinearity and was addressed above in Chapter 3. Multi-collinearity may not affect the significance of the whole model however, it inflates estimated coefficients making them unreliable.

Specifications of theapplication of thesecond model

The model appeared to be significant at a 99% confidence level.An ANOVA test of the overall model shows its significance. Table 10 reports an R Square value of 0.304 representing a proportionate reduction in the error. Therefore chosen independent variables predict more than 30% of variations of the dependent variable.

Table 10– Multiple linear regression summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.552	.304	.299	.87495

Table 11 - ANOVA test of the multiple linear regression

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	662.577	14	47.327	61.822	.000
	Residual	1514.981	1979	.766		
	Total	2177.558	1993			

The requirements of the second model are fulfilled. The distribution of regression residuals as well as the dependent variable is normal. Plotting the residuals versus predicted values allowed to exclude the issue of homoscedasticity and violation of linearity (Appendix E, page number 66). Therefore the model results can be included in the analysis.

Regression results of the second model

We first remind which hypotheses are tested using the model deleveloped before the start of the analysis of the results of the model execution. The multiple linear regression tests four of theeight formulated hypotheses. All of them are presented below.

H2: *The increase in physical proximity between two organizations facilitates their collaborative output.*

H4: *Systemic proximity between two organizations facilitates their collaborative output.*

H6: *Affiliation of both collaborators to the same university or 'non-university' sector facilitates their collaborative output.*

H8: *With the increasing technological proximity between two organizations their collaborative output first increases and then decreases.*

Table 12 presents the results of the estimation of the regression coefficients. We emphasize that there are several differences comparing to the previous model suggesting that factors influencing the initiation and facilitation of collaboration are different.

We see that distance does not have an influence on the intensity of established relationships suggesting an interesting finding. We came to the conclusion that in the German network once the link between collaborators is established distance ceases to play an important role in common activities. A possible explanation for this can be found in the technological development of a modern society. Technologies allow us to communicate effectively while being on a considerable distance from each other. However, in order to establish relationships researchers from different organizations have to meet in person or at least on line. We already discussed the exaggerated expectations that ICT eliminates the role of geography. In our results we can find some indirect support for this claim. Therefore, meeting in person is important to engage in mutual research. Here distance plays a crucial role.

Table 12–Multiple linear regression coefficients

Independent variables	Unstandardized Coefficients		Standardized Coefficients	t value	Sig.
	B	Std. Error	Beta		
Constant	-.404	.048		-8.425	.000
Distance	-.042	.040	-.041	-1.061	.289
Shared NUTS1	-.013	.106	-.004	-.118	.906
Shared NUTS2	.168	.194	.036	.866	.387
Shared NUTS3	.639	.207	.113	3.094	.002
Bordered NUTS1	-.022	.052	-.010	-.427	.670
Bordered NUTS2	.287	.089	.087	3.220	.001
Bordered NUTS3	.041	.186	.005	.220	.826
Mutual Information	.831	.079	.264	10.502	.000
Square of Mutual Information	.717	.153	.139	4.694	.000
Univ–Univ	-.074	.060	-.035	-1.225	.221
Nuniv–Nuniv	-.313	.105	-.082	-2.979	.003
Publication	.430	.044	.197	9.777	.000
Mutual Info*Univ	.602	.210	.095	2.865	.004
Mutual Info*Nuniv	.000	.319	.000	-.001	.999

Looking at systemic proximity we can not fully accept hypothesis 4. Hypothesis appeared to be right only for a low level of spatial aggregation. We can see that collaboration intensity boosts for the parties located in the same NUTS3 region. It also increases if organizations are in the bordering NUTS2 regions. Therefore, we conclude that regional characteristics play role in collaboration in Germany only at the low level of aggregation and only after the link between collaborators already exists.Organizations form a small regional innovation systems with existing research partnersexploiting the benefits of sharing a common 'social capital' (Cooke & et.al., 1997). However, when it comes to the search of new connections they do not prioritize systemically local parties but consider them equally with systemically distant parties as witnessed by our first model.

We reject hypothesis 6. Universities do not show the influence on the intensity of collaboration. But we can spot the negative affect of extra-university sector on collaboration intensity with a significance interval of 5%. We see here that productivity of extra-university sector collaborations is lower than others. Some reasons for that could be the aforementioned misalighment of objectives, stereotypes of researchers, or lack of an interface management(Heinze & Kuhlmann, 2008). Additionally, the research of these organizations may not be reflected in publications, but in patents and non-disclosure R&D projects.

An interesting result is obtained for the technological proximity. We see that mutual knowledge has only positive effect on productivity of collaborative relationships. It means that the more overlap in knowledge base organizations have the more productive is their collaborative output.At a first glance it looks like our result is contradictory to the theories about mutual knowledge. We find explanations in combination with our first model. Because our second model doesn't include failed collaborations our first suggestion is that organization simply do not engage in collaboration after a certain level of shared knowledge base is exceeded. However, that is only what our sample shows. The second possible explanation might be in the choice of the sample size and limits of our study. 100 organizations probably is not enough to capture the negative effect of mutual knowledge on collaboration intensity. In order to obtain a more comprehensive explanation a broader study of German network is needed which leaves the scope for further research. Thus we reject hypothesis 8.

Top 100 German organizations actively involve others in their research projects. The effect was shown in the first model and is repeated in the second. Organizations not only look for new research partners but actively involve third parties to participate in existing collaborations. Only half of the increase in publication productivity is reflected inestablished collaboration which is seen from positive coefficient for *publication* variable. Other half is shared with third parties.

Thus, we can make the following conclusions. The role of physical proximity on the amount of joint research projects between two organizations is eliminated in Germany. Systemic proximity is important only at the low level of spatial aggregation. 'Non-universities' are less productive when collaborating with its peers than with universities. The more overlap in knowledge base two organizations have the more productive is their research. Finally, third parties are actively involved in research projects.

6. German nanotechnology network as a part of European network

The following chapter presents the third regression model.The third model combines the approaches of the previous two. Its aim is to allow the comparison of the results with the work of Cunningham and Werker on European nanotechnology(Cunningham & Werker, 2012) and to provide a verification for the first two models. First, the chapter explains the dependent variable and presents the equation of the model. Afterwards it shows the fulfillment of its requirements. Finally, the results of the regression are analysed.

Dependent variable

The concept of collaboration is represented by the dependent variable. The variable was calculated using the fractional publication and is equal to the total amount of publications between two collaborators. If the number of publications between two organizations is zero than the variable is also zero indicating a failed collaboration.We rounded it up to the nearest integer due to the specifications of the negative binomial regression andall 4950 possible collaborations were taken into account.

Model equation

A measure of collaboration through co-authorship implies that the dependent variable will be fractional and hence can be expressed as a count variable with negative binomial probability distribution.The model is specified as follows:

$$g(Y) = \ln(Y) = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \Rightarrow$$
$$\Rightarrow \quad Y = e^{\alpha+\beta_1 X_1+\beta_2 X_2+\beta_3 X_3} = e^{x'\beta} \quad \left(x' = \begin{bmatrix} 1 & X_1 & X_2 & X_3 \end{bmatrix}\right)$$

Here Y is an integer dependent variable, α is an intercept, β is coefficient of independent variable and X_i are independent variables. The estimation of the parameters is made using the maximum likelihood technique.

Specifications of the application of thethird model

In the final negative binomial model the Omnibus test shows the level of significance equal to 0.01. The requirements of the regression are not violated. The dependent variable has a negative binomial distribution.

Negative binomial regression was chosen for the following reasons. The foundational building block in negative binomial regression framework is the Poisson regression model (Greene, 2010). However, the latter has a restriction on the distribution of the observed counts. The variance of a random variable needs to equal its mean. Because of that limitation we employ more generalized negative binomial regression. Here the variance is expected to depend on the rate of the model. In particular, the following expression for the variance is applied:

$$Var[y_i / x_i] = \lambda_i [1 + k\lambda_i]$$

An obtained data showed a strong support in favor of negative binomial regression. The likelihood ratio Chi-square for the Poisson model exceeded the value for the negative binomial model more than two times. A significant LaGrange Multiplier test (Appendix E, page 67) indentifies the over dispersion of the data, meaning that conditional variance exceeds the conditional mean. The test results can be seen in Appendix E. This fact confirms the choice of a negative binomial model instead of Poisson regression in modeling the data.

Regression results of the third model

Before presenting the results of the model we first explain differences between our model and the one constructed by Cunningham and Werker for the European nanotechnology network (Cunningham & Werker, 2012). The underlying assumptions of models are not the same. Firstly, the data time span of our model is different. If authors considered the data from 2008 to 2009, we look at publications from 2010 to 2012. Second most important difference lies in the query executed. We relied on its new version while authors used its old version. Although it is claimed that both versions of query are consistent in their findings it is important to keep this difference in mind while comparing the results of models (Arora & et.al., 2013). The models constructed are similar but not identical. The difference lies in independent variables and their interaction terms. In particular, our model didn't include the interaction term between physical and systemic proximity. Because it appeared to be insignificant and worsened model's goodness of fit. Due to specifications of the sample our network represents only academic collaboration. We coded organizational proximity differently, suggesting separation by university criteria while authors separated organizations by their academic and non-academic background.Additionally we didn't include two control variables namely total number of citations of two collaborators and total number of partnerships, due to their insignificance and high multi-collinearity with the data on total number of publications of two collaborators. However, there are more similarities than differences in both models and it is worthwhile compare their results. The operationalisation of geographical proximity and technological proximities was done in the same way. Authors also calculated collaboration intensity using fractional publication. Additionally the interaction term between organizational and technological variables have the same meaning in both models.

Thus, the third model does not focuses on formulated hypotheses of our research but instead takes a more broader view on collaboration. The model does not specify different stages of collaboration and predicts its establishment as well as intensity. A dependent variable includes all possible 4950 links between 100 organizations. We describe our results in the following order. First we compare geographical (physical and systemic) proximity, than organizational proximity and finally technological proximity.Table 13 presents all estimated coefficients with their standard errors and significance level.

Table 13 - Negative binomial regression coefficients

Independent variables	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
Intercept	-.067	.0605	-.186	.051	1.242	1	.265	.935	.830	1.052
Distance	-.161	.0508	-.260	-.061	10.033	1	.002	.851	.771	.940
Shared NUTS1	.149	.1362	-.118	.416	1.190	1	.275	1.160	.888	1.515
Shared NUTS2	.000	.2622	-.514	.514	.000	1	.999	1.000	.598	1.672
Shared NUTS3	1.037	.2735	.501	1.573	14.371	1	.000	2.820	1.650	4.820
Bordered NUTS1	.012	.0687	-.122	.147	.032	1	.858	1.012	.885	1.158
Bordered NUTS2	.291	.1123	.071	.511	6.708	1	.010	1.338	1.073	1.667
Bordered NUTS3	.084	.2443	-.395	.563	.119	1	.730	1.088	.674	1.756
Mutual Information	2.614	.1706	2.279	2.948	234.777	1	.000	13.64	9.770	19.06
Sqr Mutual Information	-4.483	.2125	-4.899	-4.066	445.083	1	.000	.011	.007	.017
Univ - Univ	-.080	.0550	-.188	.028	2.097	1	.148	.923	.829	1.029
Nuniv - NUniv	-.315	.1021	-.516	-.115	9.540	1	.002	.729	.597	.891
Publication	1.302	.0545	1.196	1.409	571.014	1	.000	3.678	3.306	4.093
Mutual Info*Univ	-.330	.2069	-.736	.076	2.542	1	.111	.719	.479	1.079
Mutual Info *NUniv	.282	.3756	-.454	1.018	.564	1	.453	1.326	.635	2.768

Our results on geographical proximity only partly overlap with the same ones obtained for the whole Europe. On the European scale a strong positive effect of physical and systemic proximity on collaboration was shown (Cunningham & Werker, 2012) while our results showed such effect only for physical proximity and low aggregational level of systemic proximity. In particular, inter-organizational collaboration benefits only if organizations are located in the same NUTS3 region. We offer the following explanation for this difference. At the international European level the coordination of nanotechnology activities is performed less than on the country level. At the country level parties have more possibilities to meet each other through various nanotechnology initiatives, conferences and projects. An example of these can be a German innovation initiative in nanotechnology (Rieke & Bachmann, 2004)(Federal Ministry of Education and Research, 2012). Such policies are supported by the government and are aimed to integrate nanotechnology partners in order to achieve success in the marketplace and innovation. Having these channels of communication the role of physical and systemic proximities is decreased if not eliminated. However, at the European level where there is a considerably less integration this role of distance is still important.

Looking at the organizational proximity we confirm our results about extra-university sector and the lack of collaboration in it. At the European level organizational proximity showed no significant influence on collaboration intensity however mediated the influence of technological proximity. In Germany we can see no mediation effect and no organizational effect except for ‘non-university sector’. The reason for such a result was mentioned in the previous two models. However, we can not compare the results, because our network representa only academic collaborations while European network represents both academic and non-academic collaborations.

The most strong confirmation is seen for the technological proximity. The character of the dependence of collaboration on the shared knowledge base in Germany is the same as the one observed in European nanotechnology. The positive effect of technological proximity can be seen in the positive significant coefficient for mutual information. Its negative effect is spotted in the negative significant coefficient for the square of mutual information. Besides, we compared our coefficients of the third model with coefficients for European network obtained by Werker and Cunningham. It turned out that the effect of technological proximity in German network has higher amplitude. A shared knowledge base is more valued by players in German nanotechnology than European players comparing to other dimensions of proximity.Figure **7**shows the observed inverted U-shaped dependence of collaboration intensity from technological proximity. It can be clearlyseenthat some optimum level of proximity exists at which cooperation between parties peaks. Too close and too distant organizations have low chances to engage in collaboration. The result confirms conclusions of Boschma(Boschma, 2005), Nooteboom(Nooteboom, 2007) as well as Werker and Cunningham (Cunningham & Werker, 2012)showing that for better functionality proximity requires some but not too greatdistance between authors.

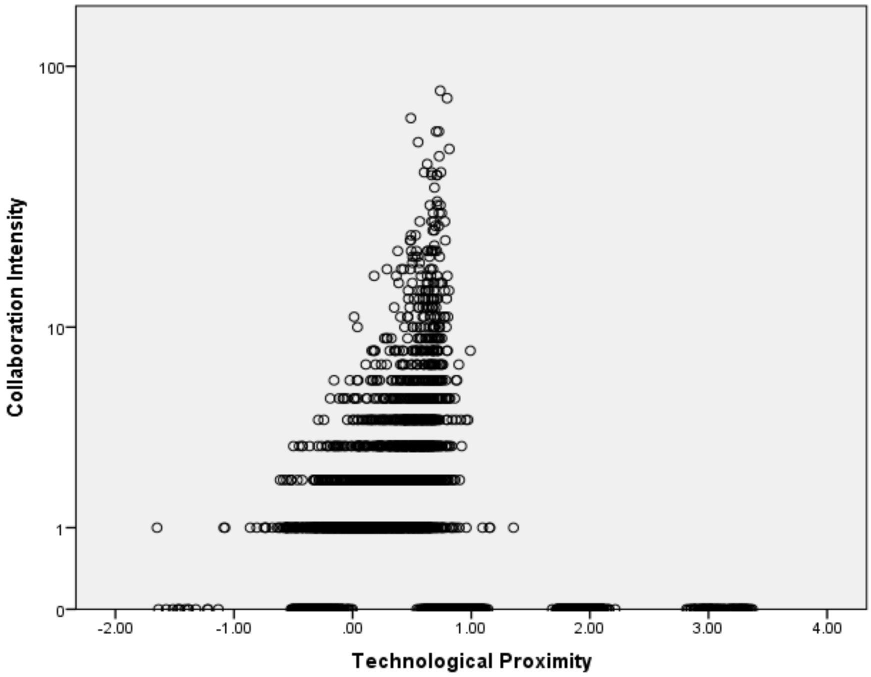


Figure 7 – Collaborations intensity (log scale) depending on the technological proximity

Combining these results with the results of the previous two models we observe a sharp drop in collaboration intensity after a certain level. The first model showed that after a certain limit in shared knowledge base German organanizations stop to collaborate. The second model omitting failed collaborations showed that the more overlap in mutual knowledge organizations have the more productive is their joint work. In the figure we see a combination of these results captured by the third constructed model.

Thus we conclude that factors that play role in European nanotechnology network are only partly present in German network. In particular, we identified three common relationships.Mutual knowledge base has an inverted U-shaped relationship with the intensity of collaboration between parties in both networks. Physical distance is negatively influencing collaborative activities. Organizations benefit from location in the same region, but only on the low level of spatial aggregation.

7. Limitations and assumptions

The following section discusses limitations and assumptions of the proposed methodology as well as their threats to the results of the research. Most of the limitations and assumptions of the methodology follow from the choice of the bibliometric analysis of the network.

Firstly, the study doesn’t include organizations that do not publish in nanotechnology. This limitation could be the reason for the poor representation of firms in the studied sample. Publications mainly reflect the production of new knowledge which is often not applicable. While firms especially small and medium enterprises are usually interested in commercial knowledge application. Additionally due to the market competition firms are not interested in sharing their developments and therefore do not contribute to the process of knowledge exchange and collaboration.

Secondly, our network is constructed without taking into account other links between organizations such as research partnerships or joint projects that do not result in publications. Grasping these agreements could help to better explain a small role played by geography in German nanotechnology collaboration. However, an inclusion of these projects would not be feasible in the time frame of the master thesis. If these collaborations do not result in publication then the added value of such projects is only shared among its partners and will never be shared with others. Thus it will not contribute to the process of knowledge exchange and collaboration.

Third limitation follows from the choice of the study of inter-organizational collaboration. Our research is omitting intra-organizational collaboration. This choice was supported by the interest of the author of this master thesis.

Fourthly, anassumption follows from the choice of the measure of collaboration through co-authorship. A contribution of each partner to the research is assumed to be equal. This issue could be solved qualitatively, conducting the interviews with authors. However, it is not feasible and eliminates all the advantages of quantitative research. Also, in thebibliometric study we may face a possible mismatch between the provided address of an organization and the actual place where research was conducted. However,in our sample the vast majority of authors had only one affiliation. Therefore it is most probable thatauthors physically appear in those places. If so our predictions concerning geographical proximity must be right.

Fifthly, a number of assumptions arise from the reliance of the research on nanotechnology query. Identified through publications organizations may not be involved in nanotechnology activity. However, these threat is avoided by the sampling procedure, which selected only 100 most productive organizations. Each of them waschecked for the existence of nanotechnology research while filling the information necessary for the regression analysis. Moreover, another assumption concerns whether these query captures all relevant nanotechnology papers. This weakness of the methodology is embedded in the query itself and can not be addressed in this project.

A final assumption was made whileoperationalising the concept ofphysical proximity. We assumed the location of every organization to be at its main office. If the assumption does not hold than it imposes threat to the results of physical proximity. However, all 100 organizations had there main research facilities near the addresses used in our study eliminating the threat.

8. Conclusion

The master thesis reveals the affect of different proximity dimensions on collaboration in German nanotechnology network. We proposed possible explanations of the discovered relationships basing on the works of other researchers.

In the introduction we mentioned that nanotechnology drives innovation breaking through the boundaries of existing technologies. Interdisciplinarity and pervasiveness allow it to underpin the vast amount of products being a general purpose technology of our time. NST is rapidly developing evidenced by an increasing number of publications (Youtie, Shapira, & Porter, 2008), diffusion of new sciences into it (Islam & Miyazaki, 2009) and rising market share of products produced withit(Roco & et.al., 2011).It leads to the growth of nanoscience networks as well as communication and cooperation of their actors.The source ofnanotechnology innovation is laying in these scientific collaboration networks. As indicated by Heinze and Kuhlmann there is little known about the factors that influence the capability of public research systems to connect knowledge flows and competences across institutional and organizational boundaries (Heinze & Kuhlmann, 2008). These flows and competences can be connected through collaboration.

We studied research collaboration in the literature together with its relation to other concepts. We showed the complexity of its definition as well as its several forms. In our work we focused on inter-organizational collaboration excluding interactions within organizations. It was found out that collaboration is crucial for knowledge processes in scientific networks. In particular, different forms of it drive knowledge creation and application. We also demonstrated findings of other researchers regarding the influence of proximity on collaboration in different nanotechnology networks. The research question addressed in our work was formulated as follows.

How do different dimensions of proximity affect collaboration and knowledge transfer in the German nanotechnology network?

We emphasize the relevance of this research because of the following three factors:

- For the first time the relationship between these concepts is studied in Germany;
- It adds to the understanding of the processes happening in nanotechnology networks;
- It may help to formulate a correct policy and management decisions to increase collaborative output of nanotechnology organizations.

We used a quantitative bibliometricanalysis to answer our research question. This method is relatively inexpensive and verifiable allowing for a large sample size(Katz & Martin, 1997). Collaboration was measured based on a multiple authorship. We considered a paper written by authors from different organizations as a collaboration. We were able to identify 100 most productive organizations calculating the amount oftheir publications. These parties served as a sample for our research. A constructed network of these organizations mainly reflects academic collaborations due to the poor representation of firms in it. It turned out that the leading organizations in German nanotechnology haven't change much for the last four years and extra-university sector continues to play a very important role in it.

In order to answer the research question we focused on thefour dimensions of proximity: physical, systemic, organizational and technological. Physical proximity represents the spatial distance between parties. Systemic proximity accounts for administrative territorial aggregation presented by the European NUTS regions. Organizational proximity is an affiliation of parties either to a university sector or to a'non-university' sector. Finally,our measure of technological proximity reflects a shared knowledge base between two collaborators. The knowledge base wasmeasured using science categories of the Web of Knowledge database. We claim that in Germany there is a difference between factors that initiate and facilitate collaboration in nanotechnology. The data on all collaborations were loaded in three regression models. Distinquishing between establishing and facilitation of collaboration we were able to show how differently proximity affects both of these processes.

The following results were obtained for the influence of proximity on the likelihood of engaging in collaboration. Physical proximity has an influence on the way organizations chose their future collaborative partners. In particular, the longer is the distance between two organizations the less are the chances that they will start to collaborate. These results are intuitive and aligned with the findings of other scholars. However, regional innovation systems do not influence this process. While looking for a new partners organizations suppress systemic proximity and do not distinquish between different NUTS regions. Universities in Germany tend to seek new partners among other universities more than in extra-university sector where organizations work together significantly less. Additionally we were able to show that universities value a mutual knowledge base less than extra-university sector when considering new collaborations.The mutual knowledge base has an optimum level which ensures a maximum productive output of organizations.Collaborations are less productive below this level and fail to appear above it. Finally, engaging in nanotechnology research organizations tend to seek for partners for the purposes of sharing costs and resources.

The second model showed the following results for the influence of proximity on the intensity of collaboration. Unlike the first model organizations are tolerant to the physical proximity when they want to develop their relationships. We explain this results with the developed communication technologies and infrastructures in Germany. At the low level of systemic proximity in particular NUTS3 regions organizations exploit the benefits of a shared 'social capital' and intensify their joint research. Collaborations in the extra-university sector are less productive on average than other mixed and pure university collaborations. It is determined by the high specializations of extra-university sector and various prejudices of its researchers. An extent of the mutual knowledge base has strong positive correlation with collaboration productivity. The more is the overlap in the knowledge bases of organizations the more productive is their research. In addition, third parties are actively involved in the existing collaborative relationships.

In our third model we compared our results with the similar ones obtained for the European nanotechnology (Cunningham & Werker, 2012). The role of geographical proximity which was represented by the physical and systemic proximities is much less in Germany than in Europe for nanotechnology collaboration. It may be explained by a better coordination of activities and substitution of this dimension of proximity by its other forms (Boschma, 2005) such as technological. In particular the relationship between collaboration and technological proximity has an inverted U-shaped form and is similar to European. However, we observed a sharp decrease in collaborations after a certain limit while European network showed smoother fall in collaborations. No comparisons were made for theorganizational proximity due to the fact that our network included only academic collaborations while European network additionally included non-academic collaborations.

Our work leaves the scope for the further research. First of all, it is possible to analysethe constructed network using the methods of a social network analysis. Identification of central nodes and clustering zones will greatly enrich our understanding of German nanotechnology. Secondly, in order to study the initiators and facilitators of collaboration more deeply additional factors could be added to the analysis such as position of an organization in the network or number of its employees. The results may help to initiate new collaborations and facilitate existing ones. Thirdly, an obtained findings could be verified through the different interviews with the representatives of analysed organizations. Fourthly, an intra-organizational collaboration can be added to the analysis. Fifthly, our results can be complemented with the information about other countries including USA and Japan. And finally our sample had poor representation of firms. It will greatly endow our results if these parties are considered increasing the sample size and taking into account non-academic collaboration.

9. Management and policy recommendations

In the following we propose some management and policy implications deriving from the results regarding the hypotheses summarized in the previous chapter. We aimto facilitate collaboration and reduce the influence of proximity on it. In Germany government authorities recognize that there is still an unrealized potential in the transfer of knowledge(Federal Ministry of Education and Research, 2012). Collaborative links between organizations may be allocated in a more efficient way decreasing the negative effect of proximity. Parties in the network will benefit from this restructuring and increase their knowledge creation capacity.

In our work we were able to show that the role of physical and systemic proximities in German nanotechnology network is small. However, the negative effect of physical distance was spotted on the likelihood of organizations to engage in collaboration. This effect can be reduced by the development of the infrastructures and by the increase in the awareness of the parties about the each others research. It can be done through a centralized information system that could accumulate the data about research profiles of organizations.

Our findings regarding organizational proximity and collaboration suggest that a facilitation of cooperation is needed for 'non-university' sector. Max-Planck, Fraunhofer, Leibniz, and Helmholtz institutions may benefit from the knowledge exchange. These organizations specialize in different areas of research. We mentioned earlier that this may be a reason why they traditionally do not collaborate much. In order to overcome these obstacles special nanotechnology projects can be initiated with a focus on societal needs. These projects should involve institutions from different sides aligning their objectives. Mixed teams of researchers may help to eliminate stereotypes and prejudices that take place between them. It can also help to resolve the problem of lack of interface management. Germany's Ministry for Education and Research already have experience in realizing such projects. Since 2003 it funded several value-chain oriented collaborative projects that mainlytried to connect scientific community and commercial world (Zweck & et.al., 2008). This experience can be used to bring together institutions from extra-university sector.

Technological proximity has the highest marginal effect on the models developed. It can be seen by adding its variable consistently with and without other independent variables. In addition to that it also has one of the highest coefficients. Therefore it should be considered as one of the most important factors while making policy and management decisions. Organizations with the optimum level of technological proximity will create strong and productive collaborations. In order to do so they have to be introduced to each other. For this purpose existing policy initiatives that are already developed in Germany may be used. For example,*The Initiative Networks of Competence Germany*includes more than 9000 members from different technological fields (Federal Ministry of Education and Research, 2012). It will involve a cluster portal in the internet which has a central access to the various initiatives. Through such a portal organizations may be introduced to each other. Also, government authorities may approach management of universities and research centers directly. They can show which connections are potentially productive once being established.

10. Appendix A

An example of publication stored in Web of Knowledgethat was analysed in the master thesis project. A specific abbreviation was attached to every piece of data. The table below shows an example of one publication.

Abbreviation	Meaning of abbreviation	Example of a publication data
PT	*Publication type*	J
AB	*Abstract*	Amide linked pyro-pheophorbide a dimers, equipped …
AF	*Authors Full*	Nikkonen, Taru Haavikko, Raisa Helaja, Juho
AU	*Authors*	Nikkonen, T Haavikko, R Helaja, J
BP	*Begin page*	2046
C1	*Addresses*	[Nikkonen, Taru; Haavikko, Raisa; Helaja, Juho] Univ Helsinki, DeptChem, Organ Chem Lab, FIN-00014 Helsinki, Finland.
CR	*Cited references*	ABRAHAM RJ, 1991, CHLOROPHYLLS, P797 ADRIAN JC, 1991, J AM CHEM SOC, V113, P678 BOXER SG, 1976, J AM CHEM SOC, V98, P5406 BOXER SG, 1981, ISRAEL J CHEM, V21, P259 BUCKS RR, 1982, J AM CHEM SOC, V104, P340 DAVYDOV AS, 1964, SOV PHYS USP, V7, P145 DEISENHOFER J, 1985, NATURE, V318, P618 GAMBLIN SJ, 2004, SCIENCE, V303, P1838, DOI 10.1126/science.1093155 …
DE	*Author Key words*	Pyro-pheophorbide …
DI	*Digital Object Identifier*	10.1039/b819764d
DT	*Document Type*	Article
EM	*e-mail*	juho.helaja@helsinki.fi
EP	*End page*	2052
FU	*Funding [grant number]*	Academy of Finland [118586, 113317]; TEKES [40407/06]
FX	*Funding text*	Prof. JacubPsencik, Charles University, Czech Republic is praised for his contribution in the interpretation of the optical spectra.

		This work was partially supported by the Academy of Finland [JH No. 118586 and 113317] and TEKES [40407/06]. The National Centre for Scientific Computing (CSC) is acknowledged for computational resources.
GA	*IDS Number*	442UH
ID	*Key Words Plus*	DENSITY-FUNCTIONAL THEORY; ENERGY-TRANSFER; ZINC CHLORIN; BACTERIOCHLORIN; DERIVATIVES; RESOLUTION; COMPLEXES; AGGREGATE; ANTENNA
IS	*Issue*	10
J9	*Journal*	ORG BIOMOL CHEM
LA	*Language*	English
NR	*Number of cited references*	34
PA	*Publisher Address*	THOMAS GRAHAM HOUSE, SCIENCE PARK, MILTON RD, CAMBRIDGE CB4 0WF, CAMBS, ENGLAND
PD	*Publication month*	
PG	*Page*	7
PI		CAMBRIDGE
PU	*Publisher*	ROYAL SOC CHEMISTRY
PY	*Publication Year*	2009
RP	*Research place*	Helaja, J, Univ Helsinki, DeptChem, Organ Chem Lab, AI VirtasenAukio 1,POB 55, FIN-00014 Helsinki, Finland.
SC	*ISI Web of Science categories*	Chemistry, Organic
SN	*ISSN*	1477-0520
SO	*Source*	ORGANIC & BIOMOLECULAR CHEMISTRY
TC	*Times cited from web of science*	1
TI	*Topic*	Hydrogen bond driven self-assembled C-2-symmetric chlorinsyn dimers; unorthodox models for chlorophyll 'special pairs' in photosynthetic reaction centres
UT		ISI:000265865600008
VL	*Volume*	7

11. Appendix B

Appendix presents three examples of the multiple spelling of organization names in theWeb of Knowledge database.

RWTH University of Aachen	Max Planck Institute for Metal Research	Leibniz Institute for Polymer Research
Univ Aachen	Max Planck Inst Met Res	Leibniz InstPolymerfroschDredsen
RVVTH Aachen Univ	Max Planck InstMetallforsch	Leibniz InstPolymerforsch Dresden eV IPF
RTWH Aachen Univ	Max Planck Inst Ferrous Res	Ipf Leibniz InstPolymerforsch Dresden eV
RheinWestfal TH Aachen	Max Planck InstMetallforschung	Leibniz Inst Polymer Res Dresden
RheinWestfael TH Aachen Univ	Max Planck Inst Metal Forsch	Leibniz InstPolymerforsch Dresden eV
RheinWestfal TH Aachen Univ	Max Plank InstMetallforsch	Leibniz InstPolymerforsch Dresden EV
Aachen Tech Univ RWTH	Max Planck InstMmetallforsch	Leibniz InstPolymerforsch Dresden e V
Univ Aachen RWTH	Max Planck InstMetallforch	Leibniz Inst Polymer Res Dresden IPF
ITMC RWTH Aachen	Max Planck Inst r Metallforsch	Leibniz Inst Polymer Res Dresden eV
RheinWestfal TH RWTH Aachen Univ	Max Planck InstMetaforchungNeue	Leibniz Inst Polymer Forsch Dresden eV
Rhine Westphalia InstTechnol RWTH	Mat &Biosyst	Leibniz InstPolymerforch
PhysChem RWTH Aachen Univ	Max Plank InstMetallforchung	Leibniz InstPolymerforchung Dresden e V
Aachen Univ RWTH	MPI Met Forsch	Leibniz InstPolymerforschung Dresden eV
RWTH Univ	MPI MF	Leibniz Inst r Polymerforsch Dresden eV
RWTH Aachen UnivTechnol	MPI Met Res	Leibniz InstPolymerforschung Dresden
DWI eV	MPI Met Forsch	Leibniz Inst Polymer Res Dresden eV IPF
RWTH Univ Aachen eV	MPI Metallforsch	Leibniz InstPolymerforschung Dresden EV
RWTH Aachen UniveV	Max Planck Inst Intelligent Syst	Leibniz InstPolymerfroschDredsen
RWTH Aachen Univ EV	Max Planck Inst Intelligente Syst	Inst Polymer Res Dresden
RWTH Aachen EV	Max Planck Inst Intelligence Syst	Leibniz InstPolymerfosch Dresden eV
DWI RWTH Aachen Univ	Max Planck InstMetalloforsch	Dresden eV
DWI EV	MPI IntelligenteSyst	Polymerforsch Dresden eV
RWTH Aachen eV		Leiden InstPolymerRes Dresden
RWTH Aachen e VDWI eV RWTH Aachen		Leiden InstPolymerforsch Dresden eV
DWI RWTH Aachen eV		Leiden InstPolymerRes
DWI RWTH Aachen EV		Leibniz InstPolymerRes Dresden IPE
DeutschWollforschungsinst RWTH		IPF Dresden
Aachen eV		Leibniz InstPolymerRes IPF
DeutschWollforschungsinst DWI eV		Leibniz InstPolymerRes IPF Dresden
DWI eV RWTH Aachen		Leibniz InstPolymerforsch Dresden IPF
RWTH Aachen e V		Leibniz InstPolymerRes Dresden EV
RWTH Aachen UnivHosp		Leibniz InstPolymerforsch eV
UnivHosp RWTH Aachen		Leibniz InstPolymerRes eV
UnivKlinikum RWTH Aachen		Leibniz InstPolymetforsch Dresden EV
UnivKlin RWTH Aachen		
UnivHosp Aachen		
UnivKlinikum Aachen		
RheinWestfael TH Aachen Klinikum		
Aachen UnivHosp		
RWTH UnivHosp		
UnivKlinikum RWTH		
RheinWestfal TH Aachen Klinikum		
RWTH UnivHosp Aachen		
UnivHosp RWTH Aachen Univ		
UnivHosp Aachen UKA		

12. Appendix C

Number	Organization Name	Ranking coefficient	Percentage of German publications covered	University(1) Non-university(0)	NUTS1	NUTS2	NUTS3	Lattitude	Longitude
1	University of Karlsruhe TH	676.44	3.67%	1	1	12	122	49.02206	8.367
2	University of Erlangen Nurnberg	573.52	6.70%	1	2	25	252	49.59072	11.01427
3	Technical University of Munich	448.21	9.67%	1	2	21	212	48.13661	11.57709
4	University of Munich	416.56	12.62%	1	2	21	212	48.26726	11.67346
5	Helmholtz ZentrumJulich	394.98	15.32%	0	A	A2	A26	50.92242	6.36391
6	Technical University of Dresden	381.33	17.78%	1	D	D2	D21	51.02858	13.73147
7	University of Aachen	317.32	20.05%	1	A	A2	A21	50.77535	6.08389
8	University of Essen Duisburg	316.88	22.03%	1	A	A1	A12	51.45564	7.01156
9	Max Planck Institute Polymer Research	310.28	23.85%	0	B	B3	B35	49.99286	8.24725
10	Leibnitz Institute for Solid State & Materials Research	299.29	25.87%	0	D	D2	D21	51.02636	13.72511
11	Technical University of Berlin	292.66	27.91%	1	3	30	300	52.51223	13.32713
12	University of Jena	275.92	29.64%	1	G	G0	G03	50.93003	11.58963
13	University of Stuttgart	262.98	31.45%	1	1	11	111	48.76348	9.15995
14	Technical University of Darmstadt	259.66	33.15%	1	7	71	711	49.87496	8.65652
15	Westfal University of Muenster	256.91	34.85%	1	A	A3	A33	51.95825	7.59127
16	University of Wurzburg	253.72	36.82%	1	2	26	263	49.78793	9.93552
17	Ruhr University Bochum	252.14	38.67%	1	A	A5	A51	51.44441	7.26176
18	Free University of Berlin	240.21	40.16%	1	3	30	300	52.45477	13.29572
19	Johannes Gutenberg University of Mainz	230.63	41.57%	1	B	B3	B35	49.99592	8.2464
20	University of Marburg	227.96	43.31%	1	7	72	724	50.8107	8.77482
21	University of Ulm	225.84	44.98%	1	7	72	724	48.4233	9.95291
22	University of Freiburg	224.15	46.57%	1	1	13	131	47.99915	7.84813
23	Max Planck Institute Colloids & Interfaces	222.75	47.91%	0	4	42	424	52.41388	12.96915
24	University of Regensburg	212.45	49.49%	1	2	23	232	48.99945	12.0932
25	University of Bayreuth	199.37	50.68%	1	2	24	242	49.927	11.58708
26	Max Planck Institute Solid State Research	183.73	51.89%	0	1	11	111	48.74649	9.0828
27	University of Halle Wittenberg	180.01	53.29%	1	E	E0	E02	51.47197	11.97942
28	HelmholtzZentrum Berlin Mat & Energie GmbH	179.01	54.63%	0	3	30	300	52.52669	13.30549
29	Max Planck Fritz Haber Institute	172.79	55.90%	0	3	30	300	52.44859	13.28275
30	Leibniz University of Hannover	167.74	57.07%	1	9	92	929	52.3859	9.71197
31	Leibniz Institute PolymerfroschDredsen	166.62	58.26%	0	D	D2	D21	51.05041	13.73726
32	University of Heidelberg	166.35	59.41%	1	1	12	125	49.40971	8.70723
33	University of Saarland	165.66	60.60%	1	C	C0	C01	49.2547	7.03991
34	Humboldt University	164.17	61.67%	1	3	30	300	52.51841	13.39427
35	University of Kiel	161.88	62.84%	1	F	F0	F02	54.34605	10.1147
36	University of Leipzig	160.73	63.58%	1	D	D3	D32	51.33857	12.37846
37	University of Hamburg	159.39	64.60%	1	6	60	600	53.56303	9.98836
38	University of Bremen	155.33	65.52%	1	5	50	501	53.10676	8.85204
39	Max Planck Institute Microstructure Physics	143.62	66.14%	0	E	E0	E02	51.49558	11.94193
40	Max Planck Institute Met Research	122.82	67.10%	0	1	11	111	48.80594	9.32378
41	University of Konstanz	119.70	67.92%	1	1	13	138	47.68879	9.18704

Number	Organization Name	Ranking coefficient	Percentage of German publications covered	University(1) Non-university(0)	NUTS1	NUTS2	NUTS3	Lattitude	Longitude
42	University of Gottingen	117.02	68.77%	1	9	91	915	51.54219	9.93575
43	Ilmenau University of Technology	116.38	69.55%	1	G	G0	G0F	50.6838	10.93096
44	Tech University of Chemnitz	110.15	70.31%	1	D	D1	D11	50.83925	12.92748
45	Bundesanstalt Mat Forsch&Prufung Berlin	108.79	71.11%	0	3	30	300	52.44306	13.28764
46	TU Braunschweig	108.14	71.93%	1	9	91	911	52.28162	10.54596
47	Paul Drude InstituteFestkoperElektronik	105.96	72.57%	0	3	30	300	52.51271	13.39695
48	University of Tubingen	105.61	73.39%	1	1	14	142	48.52898	9.05942
49	University of Bonn	105.18	73.93%	1	A	A2	A22	50.73379	7.10233
50	Technical University of Dortmund	103.55	74.47%	1	A	A5	A52	51.4929	7.41216
51	University of Kaiserslautern	103.04	75.17%	1	B	B3	B32	49.42377	7.75511
52	University of Bielefeld	101.95	75.74%	1	A	A4	A42	52.03857	8.49564
53	University of Rostock	99.675	76.39%	1	8	80	803	54.09123	12.14479
54	Goethe University of Frankfurt	97.408	77.04%	1	7	71	712	50.11873	8.65306
55	Research Centre Dresden Rossendorf	93.940	77.66%	0	D	D2	D21	51.06209	13.90534
56	University of Cologne	89.043	78.18%	1	A	A2	A23	50.93347	6.92037
57	University of Dusseldorf	88.061	78.66%	1	A	A1	A11	51.19246	6.79318
58	University of Giessen	87.920	79.05%	1	7	72	721	50.58043	8.67713
59	University of Potsdam	81.456	79.51%	1	4	42	424	52.40244	13.01135
60	PTB Braunschweig& Berlin	77.190	79.97%	0	3	30	300	52.29653	10.46418
61	University of Paderborn	76.531	80.49%	1	A	A4	A47	51.70677	8.7711
62	GKSS Research Centre GmbH	76.057	80.89%	0	F	F0	F06	53.40715	10.42279
63	Max Planck Institute Biophysics Chemistry	69.931	81.36%	0	9	91	915	51.56214	9.97092
64	University of Augsburg	69.786	81.81%	1	2	27	271	48.33329	10.89722
65	University of Osnabruck	68.262	82.19%	1	9	94	944	52.28365	8.02548
66	Max Planck InstituteEisenforsch GmbH	66.655	82.51%	0	A	A1	A11	51.23961	6.8132
67	Clausthal University of Technology	65.485	82.89%	1	9	91	919	51.8062	10.34188
68	Institute Photon Technol Jena	63.462	83.22%	0	G	G0	G03	50.92705	11.58924
69	BASF SE	60.251	83.65%	0	B	B3	B38	49.49081	8.4148
70	Carl VonOssietzky University of Oldenburg	56.597	83.98%	1	9	94	943	53.14673	8.18312
71	Brandenburg Technical University of Cottbus	54.992	84.28%	1	4	42	422	51.76699	14.32663
72	Leibniz Institute New Materials	54.496	84.60%	0	C	C0	C01	49.23755	6.99063
73	University of Hamburg DESY	53.071	84.87%	1	6	60	600	53.57702	9.8813
74	Charite Medical University of Berlin	51.281	85.12%	1	3	30	300	52.51753	13.40667
75	Technical University of Hamburg	51.257	85.55%	1	6	60	600	53.55108	9.99368
76	LazerZentrum Hannover	49.818	85.81%	0	9	92	929	52.37589	9.73201
77	University of Magdeburg	48.758	86.10%	1	E	E0	E03	52.13914	11.64099
78	University of Kassel	46.245	86.36%	1	7	73	731	51.34435	9.85933
79	Berg University of Wuppertal	44.004	86.64%	1	A	A1	A1A	51.25621	7.15076
80	University of Siegen	43.215	86.79%	1	A	A5	A5A	50.90305	8.03078
81	Technical University of Bergakad Freiberg	42.083	87.06%	1	D	D1	D13	50.91283	13.34173
82	Forschungszentrum Karlsruhe	41.367	87.36%	0	1	12	122	49.00915	8.37994
83	Max Planck Institute Physics of Complex Systems	40.714	87.61%	0	D	D2	D21	51.02676	13.71694
84	Fraunhofer Institute Applied Solid State Physics	39.332	87.83%	0	1	13	131	48.02768	7.84523
85	Ernst Moritz Arndt University of Greifswald	37.006	88.07%	1	8	80	801	54.09495	13.37435
86	Helmholtz Centre Munich	35.823	88.15%	0	2	21	212	48.22117	11.59397

Number	Organization Name	Ranking coefficient	Percentage of German publications covered	University(1) Non-university(0)	NUTS1	NUTS2	NUTS3	Lattitude	Longitude
87	Leibniz Institute of Surface Modification	35.405	88.29%	0	D	D3	D32	51.35136	12.43161
88	Max Planck Institute Chemistry Physics Solids	32.567	88.57%	0	D	D2	D21	51.02645	13.71934
89	Max Planck Institute Kohlenforschung	32.493	88.74%	0	A	A1	A16	51.4158	6.88461
90	Max Planck Institute Science Light	29.727	88.93%	0	2	25	252	49.5614	11.00438
91	Max Planck Institute Quantum Opt	29.561	89.16%	0	2	21	21H	48.25976	11.66686
92	Max Born Institute	28.308	89.39%	0	3	30	300	52.43077	13.52833
93	Fraunhofer Institute Mech Mat	27.647	89.53%	0	1	13	131	48.03397	7.85175
94	German Cancer Research Centre	26.276	89.69%	0	1	12	125	49.41437	8.67263
95	Fraunhofer Institute for Applied Polymer Research	25.491	89.86%	0	4	42	424	52.41334	12.96758
96	GSI Helmholtz Zentrum	25.073	89.97%	0	7	71	711	49.93067	8.67993
97	Fraunhofer Institute Mat & Beam Technology	24.953	90.12%	0	D	D2	D21	51.02999	13.78242
98	Max Planck Institute Mol Cell Biology&Genetics	24.445	90.25%	0	D	D2	D21	51.05838	13.78434
99	Fraunhofer Institute Solar Energy	24.427	90.46%	0	1	13	131	48.00943	7.83455
100	Jacobs University of Bremen	23.024	90.63%	1	5	50	501	53.16705	8.65081

13. Appendix D

The following presents the descriptive statistics of the variables used in the regression models.

Dependent variable

The following table presents the descriptive statistics for binary dependent variable used in the logarithmic regression.

Dependent binary

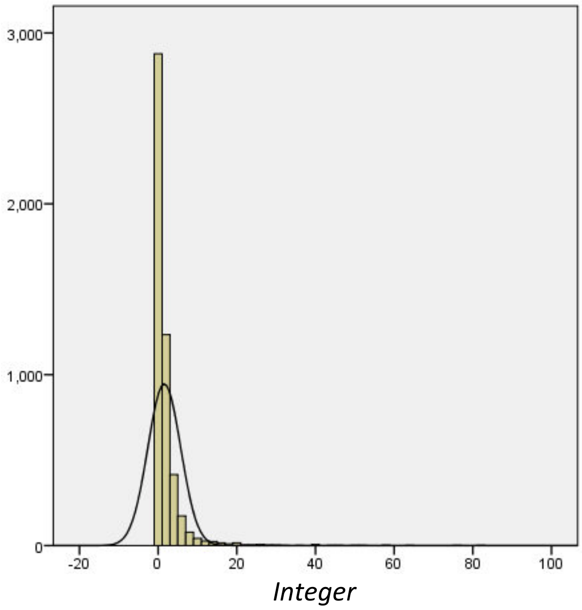
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	2878	58.1	58.1	58.1
	1.00	2072	41.9	41.9	100.0
	Total	4950	100.0	100.0	

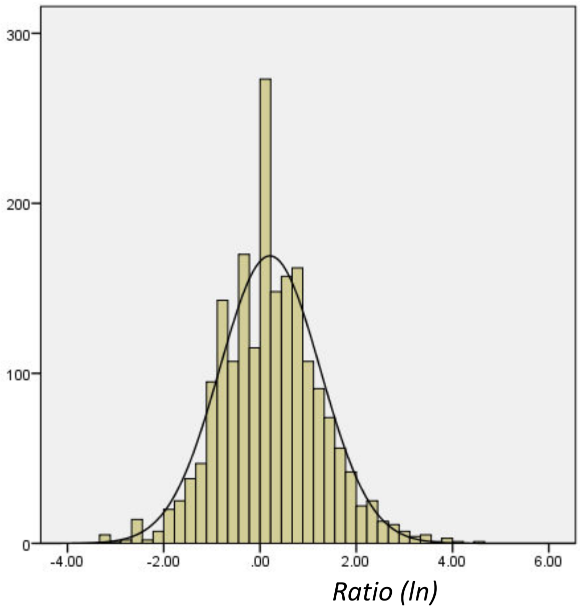
The following table presents the descriptive statistics of integer and ratio dependent variables used in negative binomial and multiple linear regressions respectively.

Statistics

		Dependent Integer	Dependent ratio
N	Valid	4950	1994
	Missing	0	2956
Mean		1.55	.2055
Std. Deviation		4.175	1.04528
Variance		17.427	1.093
Minimum		0	-3.20
Maximum		81	4.63

A corresponding histograms can be find below.





Independent variables

Two types of independent variables are presented below: categorical and continuous

Categorical variables

Shared NUTS1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	4537	91.7	91.7	91.7
	1.00	413	8.3	8.3	100.0
	Total	4950	100.0	100.0	

Shared NUTS2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	4790	96.8	96.8	96.8
	1.00	160	3.2	3.2	100.0
	Total	4950	100.0	100.0	

Shared NUTS3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	4846	97.9	97.9	97.9
	1.00	104	2.1	2.1	100.0
	Total	4950	100.0	100.0	

Bordered NUTS1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3654	73.8	73.8	73.8
	1	1296	26.2	26.2	100.0

	Total	4950	100.0	100.0	

Bordered NUTS2

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	4477	90.4	90.4	90.4
	1.00	473	9.6	9.6	100.0
	Total	4950	100.0	100.0	

Bordered NUTS3

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	.00	4902	99.0	99.0	99.0
	1.00	48	1.0	1.0	100.0
	Total	4950	100.0	100.0	

University

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	703	14.2	14.2	14.2
	1	4247	85.8	85.8	100.0
	Total	4950	100.0	100.0	

Non-university

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	1891	38.2	38.2	38.2
	1	3059	61.8	61.8	100.0
	Total	4950	100.0	100.0	

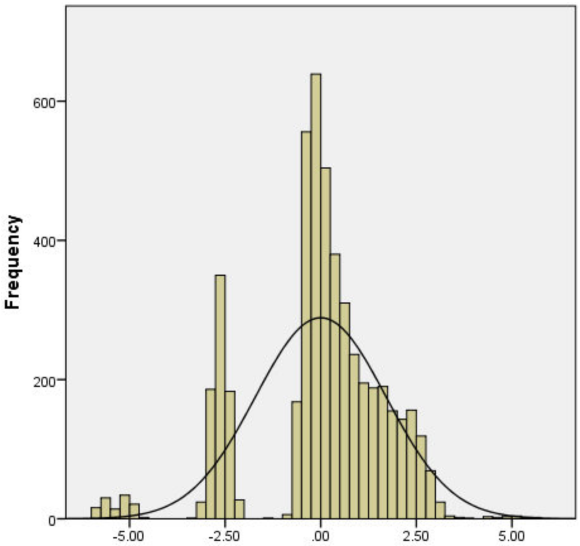
Continuous variables

A logarithmic transformation was applied to *Distance* and*Publication* variables. Additionally, all continuous variables were centered substituting or adding their means.

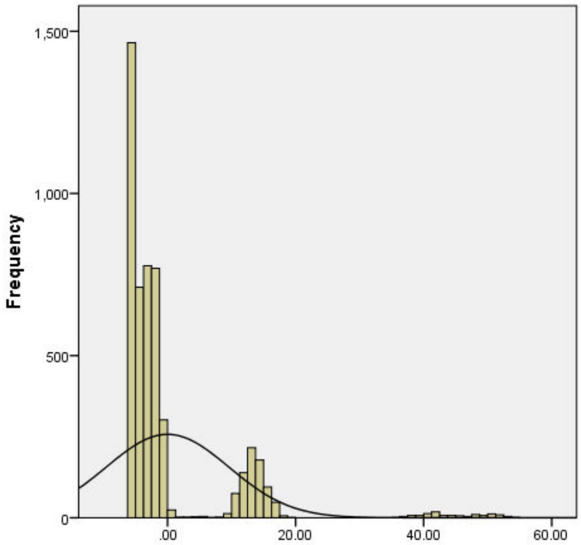
Statistics

		Mutual Information	Squared Mutual Information	Distance	Publication
N	Valid	4950	4950	4950	4950
	Missing	0	0	0	0
Mean		.0000	.0000	.0000	.0000
Median		.1019	-3.2473	.2146	-.0041
Std. Deviation		1.70931	9.59170	.85536	.54356
Variance		2.922	92.001	.732	.295
Minimum		-5.95	-6.19	-7.63	-1.40
Maximum		5.60	54.01	1.13	1.58

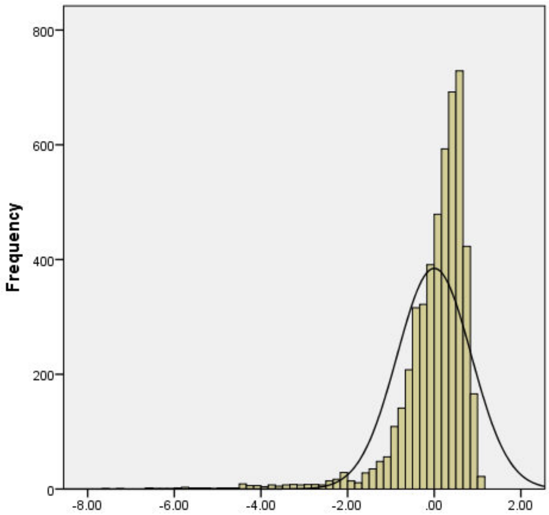
Histograms of the variables can be found below with the line showing normal distribution.



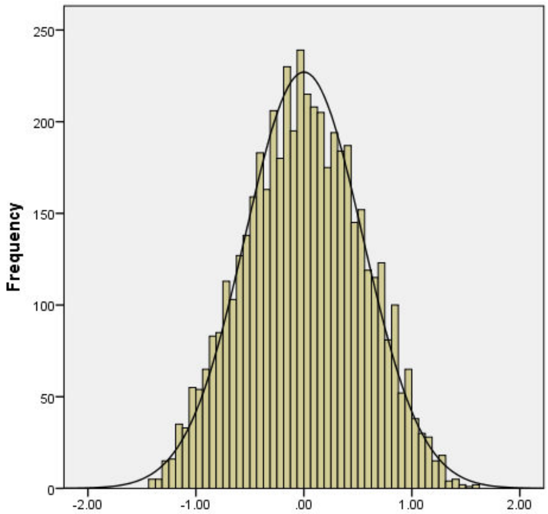
Mutual Information



Square of Mutual Information



Logarithm of distance



Total number of publications

14. Appendix E

In the following we present the output of SPSS software for all three models.

Model 1Binary logistic regression

Case Processing Summary

UnweightedCases[a]		N	Percent
Selected Cases	Included in Analysis	4950	100.0
	Missing Cases	0	.0
	Total	4950	100.0
Unselected Cases		0	.0
Total		4950	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
.00	0
1.00	1

Block 0: Beginning Block

Classification Table[a,b]

	Observed		Predicted		
			dependent		Percentage Correct
			.00	1.00	
Step 0	dependent	.00	2878	0	100.0
		1.00	2072	0	.0
	Overall Percentage				**58.1**

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-.329	.029	130.065	1	.000	.720

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	Distance	78.537	1	.000
		Shared_NUTS1	46.036	1	.000
		Shared_NUTS2	40.420	1	.000
		Shared_NUTS3	30.447	1	.000
		Bordered_NUTS1	1.807	1	.179
		Bordered_NUTS2	13.155	1	.000
		Bordered_NUTS3	6.859	1	.009
		Mutual_Information	1459.935	1	.000
		University	215.305	1	.000
		NonUniversity	112.078	1	.000
		Sqr_Mutual_Information	1002.824	1	.000
		Mutual_Inform *Univ	179.078	1	.000
		Mutual_Inform *NonUniv	156.302	1	.000
	Overall Statistics		1834.390	13	.000

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	4245.897	13	.000
	Block	4245.897	13	.000
	Model	4245.897	13	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2484.435[a]	.576	.775

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

Classification Table[a]

	Observed		Predicted		
			dependent		Percentage Correct
			.00	1.00	
Step 1	dependent	.00	2649	229	92.0
		1.00	226	1846	89.1
	Overall Percentage				**90.8**

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Distance	-.231	.120	3.694	1	.000	.794
	Shared_NUTS1	.341	.333	1.050	1	.306	1.407
	Shared_NUTS2	.853	.653	1.706	1	.192	2.347
	Shared_NUTS3	-.176	.709	.062	1	.804	.838
	Bordered_NUTS1	.282	.157	3.226	1	.072	1.326
	Bordered_NUTS2	.091	.279	.106	1	.745	1.095
	Bordered_NUTS3	-.342	.568	.362	1	.547	.710
	Mutual Information	2.777	.234	140.539	1	.000	16.076
	Univ - Univ	3.172	.375	71.580	1	.000	23.857
	Nuniv - NUniv	-1.120	.240	21.778	1	.000	.326
	Sqr_Mutual_Information	-10.197	.471	467.792	1	.000	.000
	Mutual_Inform *Univ	-6.549	.746	77.001	1	.000	.001
	Mutual_Inform *NonUniv	2.258	.583	15.029	1	.000	9.567
	Publication	2.369	.131	325.435	1	.000	10.685
	Constant	2.608	.147	314.404	1	.000	13.569

*a. Variable(s) entered on step 1: Distance, Shared_NUTS1, Shared_NUTS2, Shared_NUTS3, Bordered_NUTS1, Bordered_NUTS2, Bordered_NUTS3, Mutual_Information, Univ - Univ, Nuniv - NUniv, Sqr_Mutual_Information, Mutual_Inform *Univ, Mutual_Inform *NonUniv, Publication.*

Model 2 Multiple linear regression

Model Summary[a]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.552	.304	.299	.87495

a. Dependent Variable: Dependent

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	662.577	14	47.327	61.822	.000
	Residual	1514.981	1979	.766		
	Total	2177.558	1993			

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-.404	.048		-8.425	.000		
	Distance	-.042	.040	-.041	-1.061	.289	.230	4.342
	Shared_NUTS1	-.013	.106	-.004	-.118	.906	.333	3.000
	Shared_NUTS2	.168	.194	.036	.866	.387	.206	4.865
	Shared_NUTS3	.639	.207	.113	3.094	.002	.265	3.770
	Bordered_NUTS1	-.022	.052	-.010	-.427	.670	.709	1.411
	Bordered_NUTS2	.287	.089	.087	3.220	.001	.482	2.073
	Bordered_NUTS3	.041	.186	.005	.220	.826	.799	1.251
	MutualInformation	.831	.079	.264	10.502	.000	.558	1.791
	Sqr_MutualInformation	.717	.153	.139	4.694	.000	.399	2.507
	Univ - Univ	-.074	.060	-.035	-1.225	.221	.424	2.356
	Nuniv - NUniv	-.313	.105	-.082	-2.979	.003	.468	2.137
	Publication	.430	.044	.197	9.777	.000	.864	1.157
	MI*Univ	.602	.210	.095	2.865	.004	.319	3.135
	MI*NUniv	.000	.319	.000	-.001	.999	.448	2.234

a. Dependent Variable: Dependent

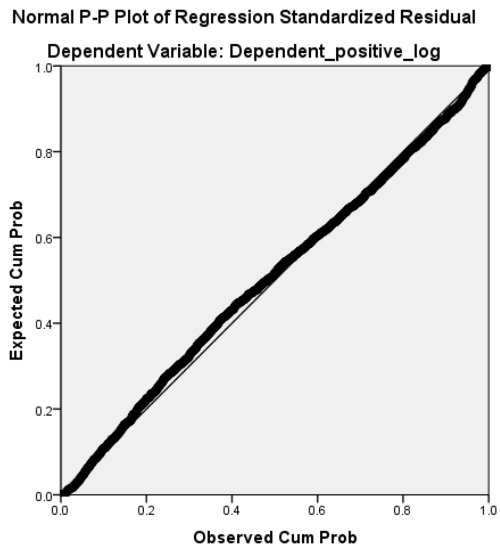
Residuals Statistics[a]

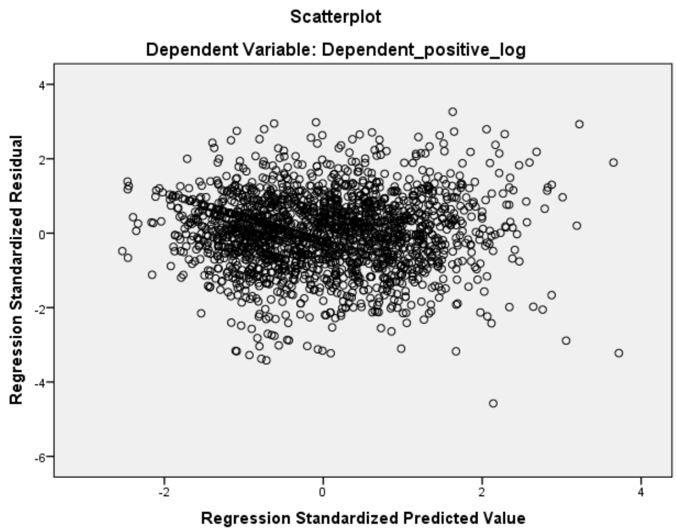
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	-1.2526	2.3504	.2055	.57659	1994
Residual	-4.00380	2.85554	.00000	.87187	1994
Std. Predicted Value	-2.529	3.720	.000	1.000	1994
Std. Residual	-4.576	3.264	.000	.996	1994

a. Dependent Variable: Dependent

Charts

The following charts plot regression residuals versus predicted value.





Model 3 Negative Binomial Regression

A Lagrange Multiplier Test showing the over dispersion.

Lagrange Multiplier Test

	z	Significance (by Alternative Hypothesis)		
		Parameter < 0	Parameter > 0	Non-directional
Ancillary Parameter[a]	7.763	1.000	.000	.000

a. Tests the null hypothesis that the negative binomial distribution ancillary parameter equals 0

Model Information

Dependent Variable	Rounded_Dependent
Probability Distribution	Negative binomial (MLE)
Link Function	Log

Case Processing Summary

	N	Percent
Included	4950	100.0%
Excluded	0	0.0%
Total	4950	100.0%

Categorical Variable Information

			N	Percent
Factor	UnivUniv	0	3059	61.8%
		1	1891	38.2%
		Total	4950	100.0%
	NUnivNUniv	0	4247	85.8%
		1	703	14.2%
		Total	4950	100.0%
	Shared_NUTS1	.00	4537	91.7%
		1.00	413	8.3%
		Total	4950	100.0%
	Shared_NUTS2	.00	4790	96.8%
		1.00	160	3.2%
		Total	4950	100.0%
	Shared_NUTS3	.00	4846	97.9%
		1.00	104	2.1%
		Total	4950	100.0%
	Bordered_NUTS1	0	3654	73.8%
		1	1296	26.2%

	Total	4950	100.0%
	.00	4477	90.4%
Bordered_NUTS2	1.00	473	9.6%
	Total	4950	100.0%
	.00	4902	99.0%
Bordered_NUTS3	1.00	48	1.0%
	Total	4950	100.0%

Goodness of model fit.

Goodness of Fit[a]

	Value	df	Value/df
Deviance	3209.750	4936	.650
Scaled Deviance	3209.750	4936	
Pearson Chi-Square	48771187.327	4936	9880.711
Scaled Pearson Chi-Square	48771187.327	4936	
Log Likelihood[b]	-4768.033		
Akaike's Information Criterion (AIC)	9564.066		
Finite Sample Corrected AIC (AICC)	9564.151		
Bayesian Information Criterion (BIC)	9655.166		
Consistent AIC (CAIC)	9669.166		

a. Information criteria are in small-is-better form.

b. The full log likelihood function is displayed and used in computing information criteria.

Significance of the model

Omnibus Test[a]

Likelihood Ratio Chi-Square	df	Sig.
2493.212	12	.000

a. Compares the fitted model against the intercept-only model.

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	6.419	1	.011
Distance	10.033	1	.002
Shared_NUTS1	1.190	1	.275

Shared_NUTS2	.000	1	.999
Shared_NUTS3	14.371	1	.000
Bordered_NUTS1	.032	1	.858
Bordered_NUTS2	6.708	1	.010
Bordered_NUTS3	.119	1	.730
Univ - Univ	2.097	1	.148
Nuniv - NUniv	9.540	1	.002
MutualInformation	234.777	1	.000
Sqr_MutualInformation	445.083	1	.000
Publication	571.014	1	.000

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Lower	Upper	Wald Chi-Square	df	Sig.		Lower	Upper
(Intercept)	-.067	.0605	-.186	.051	1.242	1	.265	.935	.830	1.052
Distance	-.161	.0508	-.260	-.061	10.033	1	.002	.851	.771	.940
[Shared_NUTS1=1.00]	.149	.1362	-.118	.416	1.190	1	.275	1.160	.888	1.515
[Shared_NUTS1=.00]	0[a]	.	.	.	.	.	.	1	.	.
[Shared_NUTS2=1.00]	.000	.2622	-.514	.514	.000	1	.999	1.000	.598	1.672
[Shared_NUTS2=.00]	0[a]	.	.	.	.	.	.	1	.	.
[Shared_NUTS3=1.00]	1.037	.2735	.501	1.573	14.371	1	.000	2.820	1.650	4.820
[Shared_NUTS3=.00]	0[a]	.	.	.	.	.	.	1	.	.
[Bordered_NUTS1=1]	.012	.0687	-.122	.147	.032	1	.858	1.012	.885	1.158
[Bordered_NUTS1=0]	0[a]	.	.	.	.	.	.	1	.	.
[Bordered_NUTS2=1.00]	.291	.1123	.071	.511	6.708	1	.010	1.338	1.073	1.667
[Bordered_NUTS2=.00]	0[a]	.	.	.	.	.	.	1	.	.
[Bordered_NUTS3=1.00]	.084	.2443	-.395	.563	.119	1	.730	1.088	.674	1.756
[Bordered_NUTS3=.00]	0[a]	.	.	.	.	.	.	1	.	.
Mutual_Information	2.614	.1706	2.279	2.948	234.777	1	.000	13.648	9.770	19.066
Sqr_Mutual_Information	-4.483	.2125	-4.899	-4.066	445.083	1	.000	.011	.007	.017
[UnivUniv=1]	-.080	.0550	-.188	.028	2.097	1	.148	.923	.829	1.029
[UnivUniv=0]	0[a]	.	.	.	.	.	.	1	.	.
[NUnivNUniv=1]	-.315	.1021	-.516	-.115	9.540	1	.002	.729	.597	.891
[NUnivNUniv=0]	0[a]	.	.	.	.	.	.	1	.	.
Publication	1.302	.0545	1.196	1.409	571.014	1	.000	3.678	3.306	4.093
MuInform*Univ	-.330	.2069	-.736	.076	2.542	1	.111	.719	.479	1.079
MuInform*NUniv	.282	.3756	-.454	1.018	.564	1	.453	1.326	.635	2.768
(Scale)	1[b]									
(Negative binomial)	1.080	.0535	.980	1.190						

a. Set to zero because this parameter is redundant.

b. Fixed at the displayed value.

15. Bibliography

Ahuja, G. (2000). The Duality of Collaboration: Inducements and Opportunities in the Formation of Interfirm Linkages. *Strategic Management Journal, 21* (3), 317-343.

Arora, S. K., & et.al. (2013). Capturing New Developments in an Emerging Technology: an Updated Search Strategy for Identifying Nanotechnology Research Outputs. *Scientometrics, 95* (1), 351-370.

Belussi, F., & Caldari, K. (2008). At the Origin of the Industrial District: Alfred Marschall and the Cambridge School. *Cambridge Journal of Economics* (33), 335-355.

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

Brenner, T., & et.al. (2011). Regional Innovation Systems, Clusters, and Knowledge Networking. *Regional Science, 90* (2), 243-249.

Cientifica Ltd. (2011, July). *Global Funding of Nanotechnologies and its Impact.* Retrieved 01 03, 2013, from http://cientifica.com/wp-content/uploads/downloads/2011/07/Global-Nanotechnology-Funding-Report-2011.pdf

Cooke, P., & et.al. (1997). Regional Innovation Systems: Institutional and Organizational Dimensions. *Research Policy, 26* (4-5), 475-491.

Cramton, C. D. (2001). The Mutual Knowledge Problem and Its Consequences for Dispersed Collaboration. *Organization Science, 12* (3), 346-371.

Cunningham, S., & Werker, C. (2012). Proximity and Collaboration in European Nanotechnology. *Regional science, 91* (4), 723-742.

Dang, Y., & et.al. (2012). Theory-Informed Design and Evaluation of an Advanced Search and Knowledge Mapping System in Nanotechnology. *Journal of Management Information Systems, 28* (4), 99-127.

Dyer, J. H., & Nobeoka, K. (2000). Creating and Managing a High-Performance Knowledge-Sharing Network: The Toyota Case. *Strategic Management Journal, 21* (3), 345-367.

European Comission. (2012, October 03). *European Comission*. (European Comission) Retrieved May 31, 2013, from Nanotechnology: http://ec.europa.eu/nanotechnology/index_en.html

Federal Ministry of Education and Research. (2012). *Federal Report on Research and Innovation 2012.* Bonn: Federal Ministry of Education and Research.

Frenken, K. (2010). The Citation Impact of Research Collaboration in Science-Based Idustries: A Spatial-Institutional Analysis. *Regional Science, 89* (2), 351-371.

Greene, W. (2010). Functional Forms for the Negative Binomial Model for Count Data. *Economics Letters, 107* (2), 313.

Grieneisen, M., & Zhang, M. (2011). Nanoscience and Nanotechnology: Evolving Definitions and Growing Footprint on the Scientific Landscape. *Small, 7* (20), 2836-2839.

Heinze, T., & Kuhlmann, S. (2008). Across Institutional Boundaries? Research Collaboration in German Public Sector Nanoscience. *Research Policy, 37* (5), 888 - 899.

Huang, C., & et.al. (2011). Nanoscience and Technology Publications and Patents: A Review of Social Science Studies and Search Strategies. *Journal of Technology Transfer, 36* (2), 145-172.

Heinze, T., & Bauer, G. (2007). Characterizing Creative Scientists in Nano-S&T: Productivity, Multidisciplinarity, and Network Brokerage in a Longitudinal Perspective. *Scientometrics, 70* (3), 811-830.

Heinze, T. (2004). Nanoscience and Nanotechnology in Europe: Analysis of publications and patent applications including comparisons with the United States. *Nanotechnology Law & Business, 1* (4).

Islam, N., & Miyazaki, K. (2009). Nanotechnology Innovation System: Understanding Hidden Dynamics of Nanoscience Fusion Trajectories. *Technological Forecasting and Social Change, 76* (1), 128-140.

Jansen, D., & et.al. (2010). Knowledge Production and the Structure of Collaboration. *Scientometrics, 83* (1), 219-241.

Kahane, B., & Mogoutov, A. (2007). Data Search Strategy for Science and Technology Emergence: A Scalable and Evolutionary Query for Nanotechnology Tracking. *Research Policy, 36* (6), 893-903.

Katz, J., & Martin, B. (1997). What is Research Collaboration? *Research Policy, 26* (1), 1-18.

Knoben, J., & Oerlemans, L. (2006). Proximity and Inter-Organizational Collaboration: A Literature Review. *International Journal of Management Reviews, 8* (5), 71-89.

Kostoff, R., & et.al. (2007). Global Nanotechnology Research Literature Overview. *Technological Forecasting and Social Change, 74* (9), 1733-1747.

Lavie, D., & Drori, I. (2012). Collaborating for Knowledge Creation and Application: The Case of Nanotechnology Research Programs. *Organization Science, 23* (3), 704-724.

Maggioni, A., & et.al. (2007). Space Versus Networks in the Geography of Innovation: A European Analysis. *Regional Science, 86* (3), 471-493.

Morgan, K. (2004). The Exaggerated Death of Geography: Learning, Proximity and Territorial Innovation Systems. *Journal of Economic Geography, 4* (1), 3-21.

Newman, M. (2001). The Structure of Scientific Collaboration Networks. *Proceedings of the National Academy of Science, 98* (2), 404-409.

Nooteboom, B. (2007). Optimal Cognitive Distance and Absorptive Capacity. *Research Policy, 36* (7), 1016-1034.

O'Brien, R. (2007). A Caution Regarding Rules of Thumb for Variance Inflatio Factors. *Quality & Quantity, 41* (5), 673-690.

Onel, S., & et.al. (2011). The Structure and Analysis of Nanotechnology Co-author and Citation Networks. *Scientometrics, 89* (1), 119-138.

Pandza, K., & et.al. (2011). Collaborative Diversity in a Nanotechnology Innovation System: Evidence from the EU Framework Programme. *Technovation, 31* (9), 476-489.

Petruzzelli, A. M. (2008). Proximity and Knowledge Gatekeepers: the Case of the Polytechnic University of Turin. *Journal of Knowledge Management, 12* (5), 34-51.

Porter, A. L., & et.al. (2008). Refining Search Terms for Nanotechnology. *Journal of Nanopartical Research, 10* (5), 715-728.

Porter, M. (1998). Clusters and the New Economics of Competition. *Harvard Business Review, 76* (6), 77.

Rieke, V., & Bachmann, G. (2004). German Innovation Initiative for Nanotechnology. *Journal of Nanoparticle Research, 6* (5), 435-446.

Rivkin, J., & et.al. (2006). Complexity, Networks and Knowledge Flow. *Research Policy, 35*, 994-1017.

Roco, M., & et.al. (2011, March). Nanotechnology Research Directions for Societal Needs in 2020: Summary of International Study. *Journal of Nanoparticle Research* , 897-919.

Ruhnau, B. (2000). Eigenvector-Centrality — a Node-Centrality? *Social Networks, 22* (4), 357-365.

Salerno, M., & et.al. (2008). Designing Foresight Studies for Nanoscience and Nanotechnology (NST) Future Developments. *Technological Forecasting and Social Change, 75* (8), 1202-1223.

Shannon, C. (1948). A Mathematical Theory of Communication. *Bell Systems Technology Journal, 27*, 379-423,623-656.

Smith-Doerr, & et.al. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quaterly, 41* (1), 116-145.

The European Commission. (2009). *Communication from the commission to the council, the european parliament and the european economic and social committee.* Brussels: Commission of the european communities.

The European Commission. (2012, September 11). *Eurostat*. Retrieved May 09, 2013, from Nomenclature of territorial units for statistics: http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction

Thomson Reuters. (2012, 05). *The Thomson reuters journal selection process*. (THOMSON REUTERS) Retrieved 01 17, 2013, from THOMSON REUTERS: http://thomsonreuters.com/products_services/science/free/essays/journal_selection_process/

Tomasello, M., & et.al. (2005). Understanding and Sharing Intentions: The origins of Cultural Cognition. *Behavioral and Brain Sciences, 28* (05), 675.

Watts, D., & Strogatz, S. (1998). Collective Dynamics of 'small world' Networks. *Nature, 393* (6684), 440-442.

Werker, C., & Cunningham, S. (2011). Policy and Concentration of Activities: The Case of Dutch Nanotechnology. *Atlanta Conference on Science and Innovation Policy.* Delft.

Youtie, J., Shapira, P., & Porter, A. (2008). Nanotechnology Publications and Citations by Leading Countries and Blocs. *Journal of Nanoparticle Research, 10* (6), 981-986.

Zweck, A., & et.al. (2008). Nanotechnology in Germany: from Forecasting to Technological Assessment to Sustainability Studies. *Journal of clenear production, 16* (8-9), 977-987.