Explaining the Evolution of Domestic Nanotechnology Companies

- Survival Patterns of a Young and Emerging Technology -

Zur Erlangung des akademischen Grades eines
Doktors der Wirtschaftswissenschaften

(Dr. rer. pol.)

von der Fakultät für Wirtschaftswissenschaften
des Karlsruher Institut(s) für Technologie (KIT)

genehmigte
DISSERTATION
von

Dipl. Wi.-Ing. Melanie Funken, geb. Vrohlings

Tag der mündlichen Prüfung: 16. Februar 2012
Referent: Prof. Dr. Kay Mitusch
Korreferentin: Prof. Dr. Ingrid Ott

2011 Karlsruhe

“There's Plenty of Room at the Bottom”
Richard P. Feynman, 1959
(FEYNMAN 1992)

Danksagung

Bedanken möchte ich mich bei allen, die mich bei der Anfertigung meiner Doktorarbeit unterstützt haben.

Hervorheben möchte ich insbesondere Herrn Prof. Dr. K. Mitusch, Frau Prof. Dr. I. Ott und Herrn Prof. Dr. H. Grupp, die es mir ermöglicht haben, meine Doktorarbeit am Institut für Wirtschaftspolitik und Wirtschaftsforschung des Karlsruher Institut für Technologie (KIT) zu verfassen und die mich wissenschaftlich stets hervorragend beraten und unterstützt haben.

Weiterhin gilt mein Dank Herrn Prof. Dr. T. Schubert, Herrn PD Dr. D. Fornahl, Herrn Dr. E. Szimba, Herrn Dr. C. A. Tran, Frau Dr. I. Haller, Frau A. Schimke und meinen lieben weiteren Kollegen, die mir mit ihrer wertvollen und konstruktiven Kritik immer sehr weitergeholfen haben und die durch ihre freundliche und unkomplizierte Art stets zu einem sehr angenehmen, positiven Arbeitsklima beigetragen haben. Nicht minder gebührt mein Dank meinen studentischen Hilfskräften für ihre ausgezeichnete und ausdauernde Unterstützung.

Besonderer Dank gebührt vor allem meiner Familie, ohne deren großartige und breite Unterstützung die Durchführung dieser Doktorarbeit nicht möglich gewesen wäre. Meinem Ehemann wie auch meinen Eltern und Geschwistern danke ich dafür, dass sie viele Stunden ihrer kostbaren Freizeit investiert haben, um mich mit Rat und Tat zu unterstützen und um mir den Rücken freizuhalten. Sie hatten stets ein offenes Ohr für mich und haben mir immer Rückhalt gegeben. Besonderer Dank gilt insbesondere auch meiner kleinen Tochter, die mich viele Male tapfer entbehrt hat und deren scheinbar grenzenlose Energie, wie auch ihr großer Forschungs- und Entdeckungsdrang mich immer wieder aufs Neue motivierten und inspirierten. Sie half und hilft mir stets, das Wesentliche nicht aus den Augen zu verlieren.

Karlsruhe, 16. März 2012

Melanie Funken

Abstract

Nanotechnology is said to be the technology of the future (see BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG 2004: 4). Due to its predicted impact, the necessity of “controlling” or “channeling” the technology grows: A proper framework needs to be provided. This book aims at shedding more light on the driving and retarding factors of nanotechnology companies. Industry-dynamic and network-theoretic methods are adapted and amended to the specifics of domestic nanotechnology firms and thus depict two new approaches to explain their evolution.

One approach focuses on the *actual* survival of companies. Data from five sources listing domestic nanotechnology companies is assembled and prepared using HOPPENSTEDT and PATSTAT (VERSION 10/ 2007) data and by conducting an extensive manual research. 354 companies are processed in the analysis. Methods of duration analysis are applied to determine the effect of pre-/ post-entry experience and technological know-how on the actual firm survival in the time between 1978 and 2009. The results of the Kaplan-Meier estimates display that the relevance of pre-entry experience cannot be assessed. In terms of post-entry experience, survival rates are lower for later entry cohorts. Technological know-how appears valuable in the long run while it seems of less relevance at the inception. However, the results of the (stratified) Cox regression display that post-entry experience alone seems to shape the observed hazard in the sample.

To better capture the characteristic of nanotechnology being a *technology*, the *technological* survival of domestic nanotechnology companies is then focused. 1284 EPO/ WIPO-nanotechnology patent applications (deriving from 382 domestic companies) are retrieved from PATSTAT (VERSION 10/ 2007) following the search strategy developed by Fraunhofer, ISI (in: NOYONS ET AL. 2003: 100-101). The patent applications are assigned to 35 technological fields according to the classification scheme suggested by SCHMOCH (2008: 9-10). A social network analysis is performed to shed some light on the relation between a company’s technological orientation and its technological survival in the time between 1978 and 2005. One finding is that companies with a higher technological overlap to other companies (core companies) tend to remain technologically active while those with a smaller technological overlap (peripheral companies) exit soon. However, as a sensitivity analysis shows, the boundary drawn between core and peripheral companies exerts a slight influence on the observed survival patterns.

In brief, there is reason to believe for experience and knowledge to at least partially explain the observed evolutionary patterns. Yet, nanotechnology is young and in a dynamic state. Political recommendations can therefore merely be of tentative nature. Long term observations need to confirm the observed patterns. The advantage of the chosen approaches thereby is that they are amendable so that they may well be adapted to the changing characteristics of the corporate domestic nanotechnology landscape. Eventually, they can be used to foster the sustainable development of domestic nanotechnology companies and thereby contribute to the country’s international attractiveness and competitiveness.

Table of contents

1 Introductory Remarks 1
1.1 Estimated impact of nanotechnology: chances and risks 1
1.2 Nanotechnology: a science-based field 3
1.3 Aim of the doctoral dissertation 6
1.4 Methodological approach 8

2 Role of Experience and Knowledge 10
2.1 Motivation and aim 10
2.2 Literature overview and choice of hypotheses 10
2.3 Database 13
2.3.1 Data ascertainment 13
2.3.2 Data collection 14
2.3.3 Data preparation: establishment of dataset 1 19
2.3.4 Data evaluation and descriptive statistics 20
2.4 Methodological approach 35
2.4.1 Kaplan-Meier estimates 35
2.4.2 Regression analysis: PH- and AFT-models 36
2.4.2.1 "Guidelines" for the adequate model choice 38
2.4.2.2 Brief outline of the Cox model 39
2.4.2.3 Final model assessment 40
2.5 In-depth analysis: significance of experience and knowledge 41
2.5.1 Nonparametric analysis: Kaplan-Meier estimates 42
2.5.2 Semiparametric analysis: Cox regression 48
2.5.2.1 Legitimation for the application of the Cox model 49
2.5.2.2 Findings of the Cox regression 51
2.5.2.3 Assessing model adequacy of the Cox model 57
2.6 Summary and conclusion 60

3 Importance of Technological Fields 64
3.1 Motivation and aim 64
3.2 Literature overview and choice of hypotheses 66
3.3 Database 69
3.3.1 Data ascertainment 69
3.3.2 Data collection 70
3.3.3 Data preparation: establishment of dataset 2 71
3.3.4 Data evaluation and descriptive statistics 71
3.4 Methodological approach 78
3.4.1 General network structure 79
3.4.2 Identification and implication of (de-)centralized networks 80
3.4.2.1 Degree-based measurement concept 82
3.4.2.2 Implications of central and decentral networks 84
3.4.3 Discovering survival patterns 87
3.5 In-depth analysis: network of technological overlap 87
3.5.1 Structural evolution 88
3.5.2 Technological centralization and decentralization 96
3.5.2.1 Centrality in the networks of technological overlap 96
3.5.2.2 Implications of the centrality/ centralization measures 100
3.5.3 Survival in the network of technological overlap 101
3.6 Summary and conclusion 116

4 Résumé ... 122
4.1 Summary of results ... 122
4.2 Political recommendations and outlook ... 127

5 Appendix ... 130
5.1 Appendix to section 1 ... 131
5.1.1 Summary of risks for mankind and for the environment ... 131
5.1.2 BMBF affirmative actions ... 132
5.2 Appendix to section 2 ... 133
5.2.1 Keyword list ... 133
5.2.2 Search strategy ... 134
5.2.3 Sources listing application fields ... 137
5.2.3.1 General ... 138
5.2.3.2 Biotechnology ... 138
5.2.4 NACE ... 138
5.2.5 Kaplan-Meier estimates: computation of mean survival times ... 141
5.2.6 Testing methods of the assumption of proportionality ... 142
5.2.7 Regression analysis: model choice ... 143
5.2.7.1 Assumption of proportionality: results of the estat phtest ... 143
5.2.7.2 Behavior of the hazard ... 144
5.2.8 Kaplan-Meier curves for the company size ... 147
5.3 Appendix to section 3 ... 148
5.3.1 Exemplary patent application ... 148
5.3.2 JAVA programs ... 149
5.3.2.1 Program 1 "build matrix" ... 149
5.3.2.2 Program 2: "multiply matrix" ... 153
5.3.3 Cooperation in the networks of technological overlap ... 158
5.3.4 WIPO IPC-Technology Concordance Table ... 162
5.3.5 Closeness- and betweenness-based measurement concepts ... 163
5.3.5.1 Closeness-based measurement concept ... 164
5.3.5.2 Betweenness-based measurement concept ... 166
5.3.6 Results of the closeness- and betweenness-based measures ... 168
5.3.6.1 Results of the closeness-based measures ... 169
5.3.6.2 Results of the betweenness-based measures ... 171

6 List of Abbreviations ... 176

7 Figures and Tables ... 177
7.1 Figures ... 177
7.2 Tables ... 178

8 Bibliography ... 179

1 INTRODUCTORY REMARKS

1.1 Estimated impact of nanotechnology: chances and risks

Nanotechnology (*colloquially* often seen as the technology concerned with particles or assembled particles on a nanoscale[1]) is considered to be the technology of the future (see BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG 2004: 4). As a technology with a cross-sectional character, "enabling technology" (see BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG 2006B: 3, 11) or "general purpose technology"[2] (see OTT AND PAPILLOUD 2007: 455), many hopes are associated with nanotechnology. From a medical point of view, great expectations exist in terms of e.g. treating cancer more precisely and efficiently or in terms of allowing for a better biocompatibility of implants (see BAUMGARTNER ET AL. 2003: 48). In connection with the information and communication technology nanotechnology is – amongst other – assumed to enable the construction of laptops which are as efficient as today's datacenters (see BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG 2006B: 12). For the environment, great innovation potential – amongst other – is seen regarding the efficient usage of resources (see HEUBACH ET AL. 2005: 34).[3] According to BHUSHAN (2005: 1) "[…] [i]t is widely felt that nanotechnology will be the next industrial revolution."

Next to the hopes and chances identified in connection with nanotechnology, fears arise concerning the potential risks that might come along with the new technology. One critical issue raised concerns the exposition of nanoparticles to the human body (see KÜHLING AND HORN 2007: 11FF). Statements of (potentially hazardous) nanoparticles overcoming biological barriers such as the blood-brain or the placental barrier repeatedly stimulate uncertainty and anxiety. Another critical issue raised concerns environmental risks (see KÜHLING AND HORN 2007: 15)[4] as well as ethical issues.[5] However, the exploration of risks is rather fragmentary.

Assuming the forecasted impact of nanotechnology is properly estimated, each nation's economic welfare is going to be increasingly reliant on the technology and its future development. Against this background, it does not occur surprising that current efforts aim at further pushing the technology forward. To give some examples, on European level *Nanoforum* (funded by the European Commission

[1] A more precise definition is discussed in section 1.2.
[2] According to OTT AND PAPILLOUD (2007: 455) a general purpose technology is characterized by pervasiveness, innovative complementarities and as having a stake in the development of societal structures.
[3] For further application opportunities see BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (2006A: 28FF).
[4] For a summary of risks for mankind and for the environment see section 5.1.1.
[5] Amongst other, STÖBER AND TÜRK (2006: 284) depict the difficulty of drawing a border between the human being and machine and furthermore point to the difficulty of defining what it is that gives a person her/ his identity.

under Framework Program 5) serves to establish links between other EU and global organizations as well as to build up networks to facilitate the exchange of information (see MORRISON 2007: 206). According to BOISSEAU (2007: 213) *Nano2Life* was launched by the European Commission under Framework Program 6 to "[…] support Europe[s] position as a competitive player and to make it a leader in nanobiotechnology transfer by merging existing European expertise and knowledge in the field of Nanobiotechnology. […]". In Germany, the BMBF[6] alone spent €27,600,000 in 1998, in 2004 it already spent €123,800,000 on research in the area of nanotechnology (see ZUKÜNFTIGE TECHNOLOGIEN CONSULTING DER VDI TECHNOLOGIEZENTRUM GMBH 2004: 32 or section 5.1.2). Thus, the amount spent on research in the area of nanotechnology was more than quadrupled in the time between 1998 and 2004.

In order to detect and exploit chances as well as to discover and avoid risks, next to gaining "technology-specific" insights, suitable framework conditions (such as a regulative framework) have to be provided.[7] Such a framework needs to allow for the (sustainable) development of nanotechnology and support a region's competitiveness at the same time. This again requires gaining an insight in factors fostering or hindering the evolution of nanotechnology. Previous works partially examine such factors. In terms of the technological evolution, HULLMANN (2001) amongst others determines which factors influence the international knowledge transfer and evaluates the impact that international knowledge transfer exerts on technological change. The author finds out that in terms of international co-publications linguistic differences and especially spatial distances hinder the international knowledge transfer to some extent, but in general exert a positive influence on the scientific and technological development of a country. International co-patents also have a positive impact on the technological development of a country – given the fact that international knowledge flows are directed into the country (outgoing knowledge flows exert a neutral influence on the technological development of a country) (see HULLMANN 2001: 251-253).

BURR ET AL. (2009) take a look at regulative and liability issues in young technology fields at the example of nanotechnology. Amongst other, the authors point out that innovation can be fostered by regulation if it is possible for the regulated company to "escape" the regulated area and if furthermore, acting in unregulated areas implies higher profits. According to the authors, innovation can also be fostered by regulation if quality regulations set incentives for intensifying R&D[8]-

[6] Bundesministerium für Bildung und Forschung (engl.: Federal Ministry of Education and Research.

[7] As can for example be illustrated by the consumer vote "BfR [Bundesinstitut für Risikobewertung]-Verbraucherkonferenz zur Nanotechnologie in Lebensmitteln, Kosmetika und Textilien" of November 20th, 2006 where, amongst other, the requirement to label groceries accordingly or to introduce admission procedures for nanoscaled substances in groceries is strengthened (see BUNDESINSTITUT FÜR RISIKOBEWERTUNG 2006: 3), a (regulative) framework is even being demanded for.

[8] Research and Development.

activities in order to fulfill regulatory requirements and to gain the permit for production or respectively, to avoid penalties. On the other hand, BURR ET AL. (2009) say that regulation hinders innovation if regulatory requirements are too strict so that R&D-activities are either slowed down, hindered, made impossible or unprofitable (see BURR ET AL. 2009: 264). As can be concluded, while HULLMANN (2001) focuses on factors influencing the technological change, BURR ET AL. (2009) specifically refer to aspects influencing the *corporate* (technological) evolution.

Heading further into the direction of explaining the evolution of companies involved in nanotechnology, HEINZE (2006) (amongst other) partially confirms his hypothesis of the *more* companies cooperate with research institutions the *higher* is their technological performance. However, according to the authors' findings, the kind of cooperation seems to be of relevance: Direct relations between companies and research institutions for example appear to lead to a higher technological performance of the companies than indirect relations. Also, the wider the spectrum of relations between companies and research institutions, the better the companies perform. The cooperation with scientifically central or internationally oriented research institutions on the other hand either has a small or no influence on a company's technological performance (see HEINZE 2006: 179-230).

Altogether, this book aims at further investigating factors driving the evolution of nanotechnology companies in Germany. The focus is on domestic companies involved in nanotechnology as they seem to increase rapidly[9]. Therefore, they are not only assumed to play a key role in terms of the *technological* development, but also to contribute largely to the nation's *competitiveness* in the area of nanotechnology. Before presenting further remarks on the aim of this book and in preparation of the analysis, it is necessary to discuss an issue which has not been covered so far: The definition of the term nanotechnology.

1.2 Nanotechnology: a science-based field

Even though previous remarks, especially the clear perception of chances and risks, might suggest else, delineating the term "nanotechnology" is difficult as a generally accepted definition of the term does not exist. The BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (2004: 6) for example uses the following definition[10]:

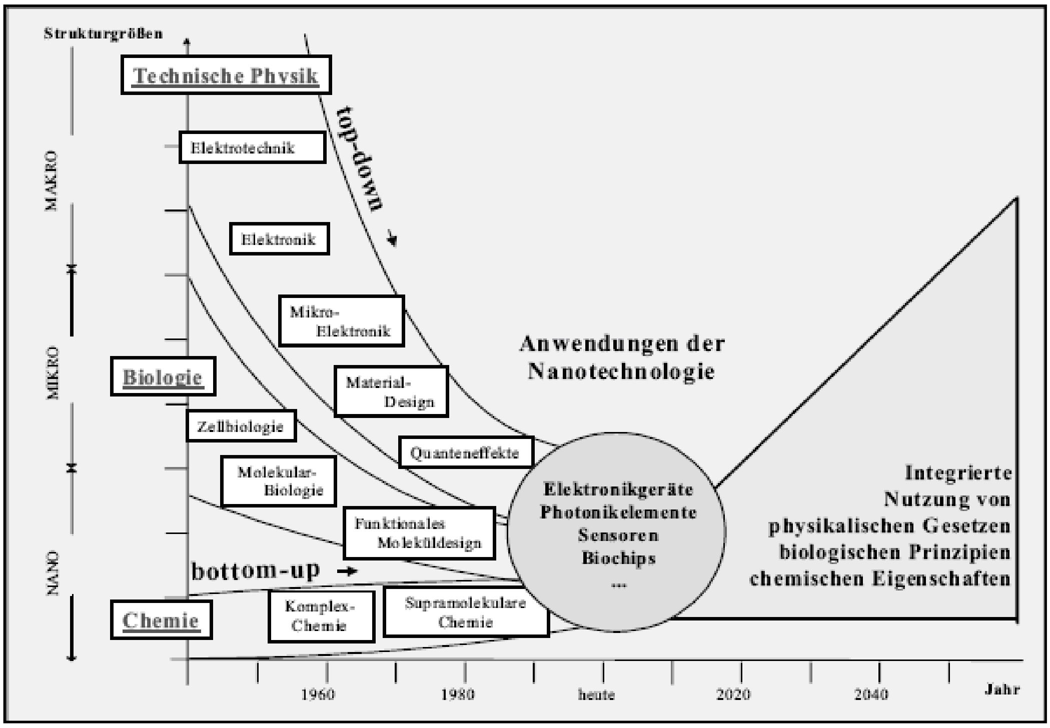
> „Nanotechnology refers to the creation, investigation and application of structures, molecular materials, internal interfaces or surfaces with at least one critical dimension or with manufacturing tolerances of (typically) less than 100 nanometres. The decisive factor is that the very nanoscale of the system components results in new functionalities and properties for improv-

[9] A glance at an overview shows that many institutions in Germany are involved in nanotechnology (for a present landscape of companies, networks and application fields etc. see VDI-TECHNOLOGIEZENTRUM: Nanomap, http://www.nano-map.de/, 7 August 2007). Timely tracking of the overview reveals that the number of institutions grows.
[10] Other definitions can be found as for example highlighted by SCHEU ET AL. (2006: 205).

ing products or developing new products and applications. These novel effects and possibilities result mainly from the ratio of surface atoms to bulk atoms and from the quantum-mechanical behaviour of the building blocks of matter."

According to ZUKÜNFTIGE TECHNOLOGIEN CONSULTING DER VDI TECHNOLOGIEZENTRUM GMBH (2004: 15-16) the absence of a widely accepted definition can be reasoned by the fact that amongst other, the boundary to microtechnology cannot easily be drawn and second, the accuracy of such boundaries is not given. Understanding this statement requires taking a closer look at the technologies history of origin depicted by Figure 1 (below).

Figure 1: General development trend and relatedness to nanotechnology.



Source: VDI TZ, in: ZUKÜNFTIGE TECHNOLOGIEN CONSULTING DER VDI TECHNOLOGIEZENTRUM GMBH (2004: 17).

The x-axis depicting time, the y-axis visualizes that nanotechnology has its roots not in one but in three different *scientific* areas simultaneously: technical physics, biology and chemistry. "Originally" each working in its "own" dimension (macro, micro or nano), over time all areas "strive" towards nanoscale, where they are able to share and exchange knowledge[11]. On atomic and molecular level, physical, chemical and biological characteristics are difficult to separate from each other (see HEINZE 2006: 109). ZUKÜNFTIGE TECHNOLOGIEN CONSULTING DER VDI TECHNOLOGIEZENTRUM GMBH (2004: 18-19) pictures three essential characteris-

[11] The integrated usage of knowledge and the continuous decrease in size seems a major source for the chances that are perceived.

tical changes occurring in the nanoworld: On the one hand these concern the quantum mechanical behavior. For instance, with decreasing particle size, the color or conductivity of a material may change. Second, due to the enlarged surface, for example the chemical reactivity might be another. Last but not least, the molecular recognition might be influenced.[12] To give an example, titanium dioxide which was proven nonhazardous on larger scales was proven toxic on nanoscale (see Hund-Rinke/ Simon 2006: 225ff cit. after FÜHR ET AL. 2006: 1FF). Altogether, working on nanoscale therefore implies that previously gained insights into all three research areas need to be reconsidered.

To exemplify the emergence of nanotechnology, the technologies' timely development is briefly sketched: According to HOLISTER (2002: 32), Richard Feynmans speech "There's Plenty of Room at the Bottom" (held in 1959) is – due to the breadth and depth of his vision and the inspiration it gave to many – often seen as year zero for nanotechnology.[13] Interestingly, the term nanotechnology does not even occur just once in Feynmans' address. At the beginning of his speech, FEYNMAN (1992: 3) says:

> "What I want to talk about is the problem of manipulating and controlling things on a small scale."

FEYNMAN (1992: 3FF) then presents detailed ideas of what could possibly be achieved in the future.[14] It is believed that in 1974 the Japanese Professor Norio Taniguchi was the first researcher to define nanotechnology[15]. His definition was the following:

> "Nano-technology' mainly consists of the processing of, separation, consolidation, and deformation of materials by one atom or by one molecule."[16]

Timely seen, *scientific* explorations seem to follow Taniguchis definition: Considering the SCI[17] (a comprehensive multidisciplinary database of scientific publications), scientific publications in the area of nanotechnology increase between 1981 and 2003; especially from 1991 on, a strong growth of publication activity is detectable (see HEINZE 2006: 108-109). Almost simultaneously, *technical* realizations assigned to nanotechnology arise. To name only few, 1974 the first patent

---

[12] These modified characteristics seem the major source for the concerns with the technology.

[13] However, HOLISTER (2002: 32) portends that "[...] there is no doubt there was nanotechnology before this and it would have developed as it has anyway, although maybe a little slower."

[14] Amongst other, he talks about putting the information contained on a page of a book on an area 1/25,000 smaller in linear scale in such manner that it is readable by an electron microscope. In his speech, he even offers a price for the person first meeting this challenge (see FEYNMAN 1992: 9). As just a couple of years later, a student manages to solve this task (see FEYNMAN 1992: 9), Feynmans' ideas do not seem outdated at all.

[15] See NANOPRODUCTS:
http://www.nanoproducts.de/index.php?mp=info&file=nanotechnologie, 21 October 2009.

[16] Cit. after NANOTECHNOLOGY RESEARCH FOUNDATION,
http://nanotechnologyresearchfoundation.org/nanohistory.html, 16 March 2009.

[17] Science Citation Index.

on a molecular electronic device is filed by Aviram and Seiden of IBM; 1981 the scanning tunneling microscope is invented by Heinrich Rohrer and Gerd Karl Binnig (see HOLISTER 2002: 32). By the help of the scanning tunneling microscope, the position of single atoms of a surface can be visualized in real space (see GOBRECHT 2006: 20). Ever since, more and more discoveries are made which are considered milestones in the area of nanotechnology (see HOLISTER 2002: 32). Next to scientific and technical explorations, manifold *products* using findings from nanotechnology are available on the market and are brought to market continuously[18]; amongst these are sunscreens and wall paint (see FÜHR ET AL. 2006: 88-93).

As a first resume it can be concluded that – though an exact definition of the term "nanotechnology" cannot be given and therefore is omitted in this work – having its roots in three scientific areas simultaneously and with a high occurrence of *scientific explorations*, *technical realizations* and *products,* nanotechnology is a strongly *science-based field*.[19] This characteristic is partially exploited in this book (see section 1.3).

1.3 Aim of the doctoral dissertation

As briefly outlined, this book aims at shedding some more light on factors driving the evolution of nanotechnology companies in Germany. Previous industry-dynamic works such as AGARWAL ET AL. (2004), KLEPPER (2002A) and THOMPSON (2003) in this respect serve as an inspiration and guiding principle for the analysis of the first three factors (pre-entry experience, post-entry experience and technological know-how).

The first factor analyzed for its impact on the evolution of domestic nanotechnology companies – or, in other words, their *actual survival*[20] – is the factor *pre-entry experience*. In brief, pre-entry experience refers to the background a firm has when entering the market. It is distinguished whether a company has pre-entry experience (it is a diversifying entrant[21], a spin-off[22] or has an experienced entrepreneur[23]) or is an inexperienced company. The hypothesis under investigation is the following:

[18] See for example NANOPRODUCTS: http://www.nanoproducts.de/, 9 March 2007. According to BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG (2006B: 11), the market potential for nanotechnology-based products is estimated up to 1.000.000.000.000 Euro in 2015.
[19] Whether in the long run, the linkage between and the development of scientific explorations, technical realizations and products is going to resemble the stylized model of market development (as described by GRUPP 1997: 33-36) remains to investigate.
[20] *Actual* survival thereby implies that a company has not exited the market due to e.g. insolvency.
[21] The company is also active in areas other than nanotechnology.
[22] The company is founded out of an existing institution.
[23] The company has a founder who previously led or owned a part of a company.

> ***HYPOTHESIS 1:*** Nanotechnology companies have a higher survival probability if they are equipped with pre-entry experience.

On the supposition that a certain age-dependency is present, the second factor analyzed is the factor *post-entry experience*. In industry-dynamic works such as CANTNER ET AL. (2006), post-entry experience refers to the experience a firm accumulates during its operation on the market. The longer a company is active on the market, the more post-entry experience it has. As explained in section 2, defining the time of market entry is problematic in case of nanotechnology. In this work, the time of market entry is therefore approximated by the foundry year of companies. The hypothesis to be tested is:

> ***HYPOTHESIS 2:*** Nanotechnology companies have a higher survival probability if they are equipped with more post-entry experience.

The remaining four hypotheses exploit the fact that nanotechnology is a science-based field. Hypothesis 3 is concerned with the presence of *knowledge* which – as in CANTNER ET AL. (2005) – is approximated by innovative activities. It is measured by the existence of patenting activities. Concretely, the following hypothesis is examined:

> ***HYPOTHESIS 3:*** Nanotechnology companies have a higher survival probability if they are equipped with technological know-how.

Next to the investigation of the impact of their "mere" presence on the *actual* survival of companies, patent applications can also be exploited else. They can be used to explain the *technological* survival of companies: Patent applications contain rich information on – for instance – their priority year, their applicant(s), but also on the patent applications *technological orientation* (the latter is depicted by IPC[24] classes which can be transferred into technological fields using concordance tables). As done in CANTNER AND GRAF (2006) this information can be used to construct networks of technological overlap for several cohorts. In the work of CANTNER AND GRAF (2006), the actors in the networks of technological overlap are institutions (referred to as innovators by CANTNER AND GRAF (2006)) and the ties between them stand for a present technological overlap between them. In other words, following CANTNER AND GRAF (2006: 466), ties or linkages emerge whenever two innovators apply for a patent in the same technological class. Based on nanotechnology patent applications, in this book, networks of technological overlap are constructed for five consecutive cohorts. In each network, the actors are domestic companies and – in line with CANTNER AND GRAF (2006: 466) – ties depict a present technological overlap between them. In each cohort, all companies following the same technological field are connected to each other (this does not imply that they cooperate). Altogether, 35 technological fields are distinguished following the WIPO IPC-Technology Concordance Table (see SCHMOCH 2008: 9-10 or section 5.3.4).

[24] International Patent Classification.

As mentioned before, the networks of technological overlap can yield another reason for survival, more specifically, *technological* survival, of companies in the area of nanotechnology. In order to test whether the technological overlap of companies accounts for firm survival, several steps have to be accomplished. Similar to CANTNER AND GRAF (2006), at first, the networks of technological overlap are examined for their basic structural attributes, including their total number of actors and their distribution to the group of entering, exiting and permanent actors. Since nanotechnology experiences an extreme growth, a strong change in the overall network organization is to expect. The hypothesis which is examined therefore is:

> ***HYPOTHESIS 4:*** The structure of the network of technological overlap changes strongly over the observed cohorts in terms of the actors which are part of the network and in terms of the intensity of ties between the actors.

Next to these basic structural circumstances, also roles of actors can be analyzed. Similar to CANTNER AND GRAF (2006) roles are distinguished into central and peripheral actors. The hypothesis under investigation is:

> ***HYPOTHESIS 5:*** In each period of time, in the networks, there are few diversified actors which – by means of social network analysis – are clearly identifiable as core actors with a high technological overlap to other actors. Consequently, there are also companies in the periphery of the network with a small technological overlap to other companies.

On the supposition that the status of actors in the networks of technological overlap is identifiable (“permanent” or “exiting” actor) and that roles are present (“core” or “peripheral” actor), lastly, both aspects are combined to explain the companies’ *technological survival* within the network. This leads to hypothesis 6:

> ***HYPOTHESIS 6:*** The majority of companies in the core of the network in cohort t remain actors of the network in cohort (t+1). The majority of companies in the periphery of the network of cohort t exit after cohort t, i.e. are not part of the network in cohort (t+1).

In summary, hypotheses 1, 2 and 3 are tested by means of a survival analysis (see section 2). Hypothesis 4, 5 and 6 are examined by means of a social network analysis (see section 3).

1.4 Methodological approach

In order to test the hypotheses and thereby to explain the evolution of nanotechnology companies, this book consists of two main sections. To ease understanding, both sections are structured similarly.

Following the introductory remarks on nanotechnology given in this section, section 2 examines the role of experience and knowledge for the *actual survival* of domestic nanotechnology companies. In preparation of the analysis, the motivation and aim of the section is outlined in section 2.1. Section 2.2 gives an overview on previous industry-dynamic studies to further elucidate and grasp the underlying research idea. Furthermore, the choice of hypotheses is explained in detail. As the performance of a survival analysis requires the presence of an extensive dataset, a comprehensive description of the sample underlying the survival analysis and the – with the sample associated – definition of nanotechnology companies is given in section 2.3. This dataset is referred to as dataset 1. In preparation of the survival analysis, the diverse methodological steps of the analysis are presented in section 2.4. Following these steps, the results of the survival analysis are presented in section 2.5. Section 2 ends with a summary and conclusion of the obtained findings (see section 2.6).

Due to the fact that various technologies influence the development of nanotechnology (and vice versa) and the pursuit of distinct technologies might themselves account for the survival of companies, section 3 focuses on the *technological survival* of firms. In section 3.1 the sections motivation and aim is presented. Section 3.2 gives a brief overview on previous studies in the area of social network analysis. Also, the hypotheses which are examined in section 3 are deduced. As in case of the survival analysis, also, performing a social network analysis requires the availability of an extensive dataset. A detailed description of the database underlying the social network analysis (referred to as dataset 2) including the (in comparison to section 2) *modified* definition of nanotechnology companies is therefore given in section 3.3. Techniques of social network analysis are introduced in section 3.4. To assess whether the “width” of the technological orientation of a company’s nanotechnology patent applications may account for a company’s (technological) survival, the respective results of the analysis are yielded in section 3.5. At last, section 3.6 ends with a summary and conclusion of the obtained results.

Finally, section 4 provides a resume. Section 4.1 summarizes the results of the previous sections. Determinants influencing the evolution of nanotechnology companies in Germany are highlighted. In section 4.2, probable guidelines for political actions are deduced and starting points for future works are given.

2 ROLE OF EXPERIENCE AND KNOWLEDGE

2.1 Motivation and aim

In a globalized world where competition amongst companies is strong, experience and knowledge are frequently assigned a central role. Experience – for instance in the sense of previously gathered market know-how – may be helpful to identify the most promising business options and finally, to render the proper short-, mid- and long-term decisions. Knowledge – for instance technological knowledge – may contribute to prevail against possible "opponents" and help strengthening the company's competitive position in the long run.

Following the notion of "experience" and "knowledge" being of crucial importance for the actual survival of companies (*actual survival* thereby meaning that a company has not exited the market due to e.g. insolvency), this section abstracts from the "single" firm perspective and instead focuses on a more general economic level. By performing a survival analysis, it is examined how experience (distinguished into pre- and post-entry experience) and knowledge (further referred to as technological know-how) may account for the evolution of nanotechnology companies in Germany.[25] Exploring these factors, this section follows the idea of a distinct strand of industry-dynamic studies (sketched in the following section); with the particularity that in this book (instead of an industry) a *technology* is in the center of attention.

2.2 Literature overview and choice of hypotheses

A large number of empirical studies analyze a distinct industry (for example the automobile industry) and track the industry incumbents in a distinct period of time. Entries into and exits from the industry are recorded and selected characteristics of firms are stored in order to give an explanation for the pattern of some firms remaining in the industry while others fail.

Such characteristics can concern the experience a company has: "Experience" may refer to the pre-entry background of a company, as outlined by KLEPPER (2002A), THOMPSON (2003), CANTNER ET AL. (2006) or AGARWAL ET AL. (2004). Usually, different "types" of experience are distinguished. For the U. S. automobile industry KLEPPER (2002A) identifies four groups: experienced firms, experienced entrepreneurs, spin-offs and inexperienced firms. He specifies diversifying firms to be experienced firms. As for experienced entrepreneurs KLEPPER (2002A: 648) argues

[25] Aside from experience and knowledge, also other factors such as political actions may influence the prosperity of individual companies or entire industries. These factors are not considered in the analysis.

> "[...] The second category includes *de novo* firms that were founded by individuals that headed, and typically owned a significant part of, a firm in another industry. [...]"

Category three (spin-offs) also contains de novo entrants, with one or more founders having experience in the specific industry. All other de novo entrants are firms in which the founders have no specific experience (inexperienced firms) (see KLEPPER 2002A: 648). Analyzing the German automobile industry CANTNER ET AL. (2006: 59) employ the same classification as KLEPPER (2002A). THOMPSON (2003) examines the shipbuilding industry and also differentiates four groups, but names and defines these slightly different than mentioned. According to THOMPSON (2003: 11-12) the first group ("experienced") includes firms with prior experience in the industry (either the shipbuilding or engine manufacturing industry). If a firm previously operated a foundry without having ventured in the engine manufacturing industry, it is labeled as group two ("foundry"). Firms of the type "miscellaneous" have prior experience in diverse fields. Finally, firms of type four have an unknown background and therefore are labeled "unknown". For the disk drive industry, AGARWAL ET AL. (2004: 508-509) identify four categories of entrants, namely spin-outs, incumbent-backed entrants, diversifying entrants and non-spin-out de novo entrants. Spin-outs are defined as follows:

> "[...] as a firm started by individuals who were employees of existing firms in the industry [...] in the year prior to the spin-outs' formation. [...]"

According to the authors, "incumbent-backed entrant" is the generic term for firms, which either are affiliated with firms in the industry, which are subsidiaries, parent-sponsored ventures or joint ventures. Furthermore, diversifying entrants are defined to be firms which existed in another industry. Finally, firms with no immediate connection to the industry and which are not diversified are called "non-spin-out de novo entrants".

Though not all authors name or define the same categories and partially examine very different industries, they do seem to agree that experience does have a profound influence on the survival of firms (see KLEPPER 2002A: 661FF, CANTNER ET AL. 2006: 57FF and THOMPSON 2003: 27). Especially spin-offs seem to be amongst the most successful entrants (see AGARWAL ET AL. 2004: 514). This is also emphasized by other studies, as in KLEPPER AND THOMPSON (2006: 11) who point out that

> "[…] in autos, disk drives, lasers, medical devices, tires and wine, the performance of spin-offs, proxied by longevity, size, scope, years to first VC[[26]] funding or pre-money valuation, is superior to other *de novo* entrants and is comparable if not superior to diversifiers from related industries. [...]"

Another frequently examined characteristic is the factor post-entry experience or rather the *time* a company enters into an industry (for instance see KLEPPER AND SIMONS 1999, HORVATH ET AL. 2000, KLEPPER 2002A, CANTNER ET AL. 2006).

[26] Venture capital.

KLEPPER AND SIMONS (1999: 2, 36FF) gather data on four industries, namely the automobile, tire, television and penicillin industry and investigate the causes for industry shakeouts. They discover that by the time of the shakeout earlier entrants have significantly lower hazard rates persisting many years thereafter. Second, in the automobile and tire industry, hazard rates seem to decline with age. In the other two industries, age does not seem to have an effect on the hazard rate. Also analyzing the U. S. automobile, but also the U. S. beer brewing and tire industry, HORVATH ET AL. (2000: 18FF) on the other hand reveal that

> "[...] the shape of exit hazard rates are qualitatively similar across the life span of an industry; regardless of when a cohort enters, the highest conditional probability of exit is in age 1-2 years. By age 5, hazard rates decline markedly to low levels and remain low for the remainder of the cohort's life. […]"

HORVATH ET AL. (2000: 18FF) point out though that there are higher hazard rates for cohorts entering late in an industry's life cycle.

Aside from both kinds of experience – pre- or post-entry experience – also *knowledge* is analyzed as a factor possibly explaining the survival of companies. Industry-dynamic works concentrate on explicit knowledge, which for example can be captured in patents or publications. CANTNER ET AL. (2005: 1FF) amongst other perform a statistical survival analysis to determine how the survival of firms is influenced by knowledge acquired through innovative activities. They find out that post-entry experience, pre-entry experience and innovative activity all are very important for the survival of firms and that it is possible for innovative activities to compensate for the lack of post- and pre-entry experience. Following the above approaches, the subsequent hypotheses are derived for this book:

> ***HYPOTHESIS 1:*** Nanotechnology companies have a higher survival probability if they are equipped with pre-entry experience.
>
> ***HYPOTHESIS 2:*** Nanotechnology companies have a higher survival probability if they are equipped with more post-entry experience.[27]
>
> ***HYPOTHESIS 3:*** Nanotechnology companies have a higher survival probability if they are equipped with technological know-how.

As mentioned, the hypotheses are tested by means of a survival analysis. Performing a survival analysis requires the availability of an extensive dataset. The sample underlying this work is described in the following sections (see sections 2.3.1 to 2.3.4).

[27] A company has *more* post-entry experience than another company if it enters earlier to the field of nanotechnology than the respective other company.

2.3 Database

2.3.1 Data ascertainment

Inspired by CANTNER ET AL. (2005: 7) and BUENSTORF (2005: 13), the list below presents a brief overview on the information which is collected for the purpose of section 2 (this section). A detailed description of each of the variables follows in section 2.3.2.

(1) Application field: NACE[28]-codes give information on the economic areas in which a company is active in.

(2) Location: "Location" depicts the companies' address and Federal State.

(3) Company size: Deduced from employment and sales data, the company size yields information on whether the company is a large company, a medium/ small sized company or a microenterprise. If – due to missing data – a category cannot be determined, the company is classified "without classification".

(4) Time frame of analysis: The survival of companies is examined in the interval of [1978, 2009]. This is the time span in which nanotechnology is often said to have strongly evolved.

(5) Time of study entry: In this book, the time of study entry is approximated by the foundry year. In accordance with the time frame of analysis, only companies with foundry years between 1978 and 2009 are admitted for analysis (the reasoning for this decision is given in later paragraphs, see section 2.3.4, paragraph (5)).

(6) Time of study exit: By 2009, a company may still be active, it may have merged with/ have been acquired by another company or it may have exited the market before (due to insolvency, liquidation or because it ceased to exist[29]). Consequently, three study "exit" cases are distinguished: "active companies", "mergers and acquisitions" and "exiting companies". The time of study exit – an essential prerequisite for the performance of a survival analysis – depends on the type of study exit: conditioned by the time frame of analysis, the reference year is set to 2009 in case of active companies, to the year of the merger or acquisition (in case of mergers or acquisitions) and to the year of exit, e.g. insolvency (in case of exiting companies).

[28] Nomenclature générale des activités économiques.

[29] These categories are "inherited" from HOPPENSTEDT, which is one data source used to complete the company data in this section.

(7) *Duration time and (right-)censoring:* The duration time is the difference between the time of study exit and study entry. Depending on the study exit type, right-censoring is either applied or not: In case of active companies, mergers or acquisitions, the respective entry is right-censored. In case of exiting companies, censoring is not applied.

(8) *Pre-entry experience:* It is distinguished whether a company has pre-entry experience (the company is either a diversifying entrant, a spin-off or has an experienced founder) or is an inexperienced company (pre-entry experience is not determinable).

(9) *Post-entry experience:* Three entry cohorts are distinguished (depending on the foundry year).

(10) *Technological know-how:* For each company in the dataset, it is recorded whether the company holds (nano) patent applications.

While in the above listing, variables “(1) application field”, “(2) location” and “(3) company size” serve to give an impression on the general characteristics of the underlying database, the remaining seven variables directly refer to the survival analysis: “(4) time frame of analysis”, “(5) time of study entry”, (6) time of study exit” and “(7) duration time and (right-)censoring” are time-related variables which need to be specified in the context of a survival analysis. “(8) pre-entry experience”, “(9) post-entry experience” and “(10) technological know-how” are the factors which are subject to investigation. To collect the data, diverse sources have been used. They are described in the following paragraphs.

2.3.2 Data collection

In industry dynamics, typically, companies “belonging” to the investigated industry are identified according to a distinct product. However, in this research work, a *technology* is in the center of attention. As such, it affects a wide range of industries and products may differ strongly from another. Together with the addressed absence of a unique, generally valid definition of the term “nanotechnology”, it is impossible to define an exact and absolutely indisputable set of companies. Therefore, in section 2 (this section) of the book, companies are identified (and defined) over *sources* listing companies (or corporate divisions)[30] which are concerned with nanotechnology. The companies furthermore need to have their headquarter or – as a minimum – at least one location in Germany. Altogether, five sources are used. The five sources build the foundation of the company database. As the subsequent paragraphs show, each of the sources has its own

[30] The data is left at this disaggregated level so that a detailed analysis is possible. For reasons of simplicity, the term “companies” is used for both in the following.

focus which on the one hand enriches the sample, on the other hand leads to certain heterogeneity.

(1) "Förderkatalog"[31] (data collection took place on 16 January 2007)

The "Förderkatalog" of the BMBF/ BMWi[32] comprises a broad set of projects and the respective organizations concerned with these projects. Therefore, before being able to extract companies, projects related to nanotechnology need to be extracted first. For this reason, the keyword "%nano%"[33] is entered in the search mask yielding the results shown in Table 1 (below).

[31] BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG/ BUNDESMINISTERIUM FÜR WIRTSCHAFT UND TECHNOLOGIE: Förderkatalog, http://oas2.ip.kp.dlr.de/foekat/foekat/foekat, 16 January 2007.
[32] Bundesministerium für Wirtschaft und Technologie (engl.: Federal Ministry of Economics and Technology).
[33] %=unlimited truncation (0 or else).

Table 1: Förderkatalog, all LP[34]-numbers and labels.[35]

LP-Ziffer	LP-Text
I210	Verfahren und Geräte für die Silizium-Mikro- und Nanoelektronik
K01010	Nanobiotechnologie
L110	Nanotechnologie - Branchenspezifische Maßnahmen
L11010	Leitinnovation NanoMobil
L11020	Leitinnovation NanoLux
L11030	Leitinnovation NanoForLife
L11050	Leitinnovation NanoChem
L111	Nanotechnologie – Prozesstechnologie
L11110	Prozesstechnik und Nanoanalytik
L112	Nanotechnologie - Interdisziplinäre Technologien
L11210	Nanobiotechnologie
L11220	Nanomedizin
L113	Nanotechnologie – Werkstoffkonzepte
L11310	Nanostrukturmaterialien
L11320	Nanokomposite
L14010	KMU incl. NanoChance
L14110	Nano-Zentren
L250	Nanotechnologie
L25010	Laterale Nanostrukturen
L25020	Nano-Optoelektronik
L25099	Sonstige Nanotechnologien und Querschnittsaktivitäten (Kompetenzzent-ren, Gutachter- und Strategiekreis)

Source: BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG/ BUNDESMINISTERIUM FÜR WIRTSCHAFT UND TECHNOLOGIE[36].

According to the database operator, the listed projects have been fostered by both institutions (BMBF/ BMWi) from the budget of nanotechnology.[37] However,

[34] Leistungsplansystematik.
[35] To avoid using false translations, the results are presented in the manner they are retrieved from the database.
[36] BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG/ BUNDESMINISTERIUM FÜR WIRTSCHAFT UND TECHNOLOGIE: Förderkatalog, http://oas2.ip.kp.dlr.de/foekat/foekat/foekat, 16 January 2007.
[37] "Nanotechnology"-projects which are *not* financed from the budget of nanotechnology are not included in the list. They are disregarded in this book as they only deliver a small contribution to the main groups other than the above mentioned ones. Additionally, finding these projects is error-prone: The project titles in the complete database would have to be searched for keywords. Merely searching for the sequence "nano" is ambiguous enough as this sequence occurs frequently in the German language; e.g. the composite "Lampe**nano**rdnung" contains the term, but does not necessarily have anything to do with

not all groups/ projects that are concerned with nanotechnology necessarily have to contain the term “nano” in their name. For this reason, in a second step, the dataset is expanded again: if in the above table the sub as well as the main group contains the term „nano“, the main group is chosen, the subgroup discarded. This leads to an enlargement of the dataset as by this procedure subgroups *not* containing the term “nano” are also included. The remaining nine groups are listed in Table 2 (below).

Table 2: Förderkatalog, main LP-numbers and labels[38].

LP-Ziffer	LP-Text
I210	Verfahren und Geräte für die Silizium-Mikro- und Nanoelektronik
K01010	Nanobiotechnologie
L110	Nanotechnologie - Branchenspezifische Maßnahmen
L111	Nanotechnologie – Prozesstechnologie
L112	Nanotechnologie - Interdisziplinäre Technologien
L113	Nanotechnologie – Werkstoffkonzepte
L14010	KMU incl. NanoChance
L14110	Nano-Zentren
L250	Nanotechnologie

Source: BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG/ BUNDESMINISTERIUM FÜR WIRTSCHAFT UND TECHNOLOGIE[39].

Having extracted the relevant projects, in a second step, companies are deduced. Other organizations are “discarded”. In total, 417 companies are retrieved by this procedure.

(2) Heinze 2006

HEINZE (2006: 114, 148FF) searches for nanotechnology patents at the EPO[40] [EPO and Euro-PCT-patents]. In total, he identifies 290 organizations (universities, other research facilities and companies) which either “directly” apply for a patent between 1991-1995 or 1996-2000 or which can be assigned a patent ap-

nanotechnology. It seems more reasonable to employ a more complicated search strategy, e.g. to use a list of keywords combinations. In a previous work a comparable list was configured by the VDI and used by HEINZE (2006) (see HEINZE (2006: 282) or section 5.2.1). As the keyword list is not up to date and since nanotechnology is an evolving and diversified technology, from a present point of view this keyword list seems obsolete: many other keywords should be included in the meantime.

[38] To avoid using false translations, the results are presented in the manner they are retrieved from the database.

[39] BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG/ BUNDESMINISTERIUM FÜR WIRTSCHAFT UND TECHNOLOGIE: Förderkatalog, http://oas2.ip.kp.dlr.de/foekat/foekat/foekat, 16 January 2007.

[40] European Patent Office.

plication after matching the inventor names *without* organizational context with the SCI. Altogether, 191 of the organizations are companies. They are considered for further investigation.

(3) Nanomap[41] (data collection took place on 7 August 2007)

"Nanomap" gives an overview on institutions concerned with nanotechnology. It is possible to filter according to the kind of institution (companies, research institutes, networks etc.), according to the application field (e.g. chemistry/ materials), or according to the technologies (e.g. nanoelectronics). Using the filter "institution", in total, 653 companies are processed for further analysis.

(4) "Nanotechnologie-Unternehmen"[42] (data collection took place on 9 March 2007)

The source lists companies which either supply or use nanotechnology. In total, 103 companies are considered for further analysis.

[41] VDI-TECHNOLOGIEZENTRUM: Nanomap, http://www.nano-map.de/, 7 August 2007.
[42] NANOTECHNOLOGIE IN HESSEN: Nanotechnologie-Unternehmen, http://www.nanoportal-hessen.de/brancheninfo/unternehmen/, 9 March 2007.

(5) Nanoproducts[43] (data collection took place on 9 March 2007)

The web page lists organizations offering products which (according to the producer) involve nanotechnology. At the time of data collection, the products are assigned to eleven main groups (see Table 3 below).

Table 3: Nanoproducts, main and subgroups.

No.	Main groups (→ Subgroups)	Companies/total
1	coatings	15/15
2	services	2/2
3	leisure time with nano	10/14
4	cosmetic	5/10
	→ raw materials	
5	literature	0/4
6	materials	17/17
7	technologies	11/11
8	surfaces	9/21
	→ anti fingerprint	
	→ anti fouling	
	→ easy to clean	
9	additives	6/7
10	nanoanalytic	17/20
11	process engineering	3/3

Source: NANOPRODUCTS[44].

The second number in the right column gives information on how many products are listed in total at the time of data collection, the first number denotes the number of products which are assigned to companies. The difference between the numbers is due to the fact that some products are developed by universities or other kinds of institutions. Altogether, 94 companies and their products are retrieved.

2.3.3 Data preparation: establishment of dataset 1

Having identified the companies, the data of the previously described five sources is merged into an EXCEL spreadsheet. After cleaning up from multiple entries, 1078 companies are left for further analysis. These are matched with data from PATSTAT (VERSION 10/ 2007)[45] and HOPPENSTEDT[46, 47]. Since not all

[43] NANOPRODUCTS: http://www.nanoproducts.de/, 9 March 2007.
[44] NANOPRODUCTS: http://www.nanoproducts.de/, 9 March 2007.
[45] The data is merged using name matching after detecting nanotechnology patent applications at the DPMA, the EPO and WIPO by the help of the search strategy developed by

companies can successfully be matched and/ or necessary preconditions (such as a too early foundry year), are not met, the sample size reduces to approximately one third of the original sample size. To close existent gaps in the database (for example in terms of the companies pre-entry experience), a very extensive and therefore *extremely* time-consuming *manual* internet research is conducted involving sources as LEXISNEXIS[48] and GENIOS[49], but also a high number of – for example – corporate web pages, press releases etc.

To secure anonymity of the study, neither the sources of manual research nor any other information revealing the companies' identity are disclosed in detail in this book. Information on the database is restricted to the statistics presented in section 2.3.4.

For altogether 354 companies the necessary data requirements for performing a survival analysis (simultaneous availability of study entry/ exit time, knowledge regarding the factors pre-/ post-entry experience and technological know-how etc.) are met. These 354 companies are further referred to as "dataset 1". They build the data basis of section 2. The subsequent paragraphs and descriptive statistics provide detailed information on dataset 1.

2.3.4 Data evaluation and descriptive statistics

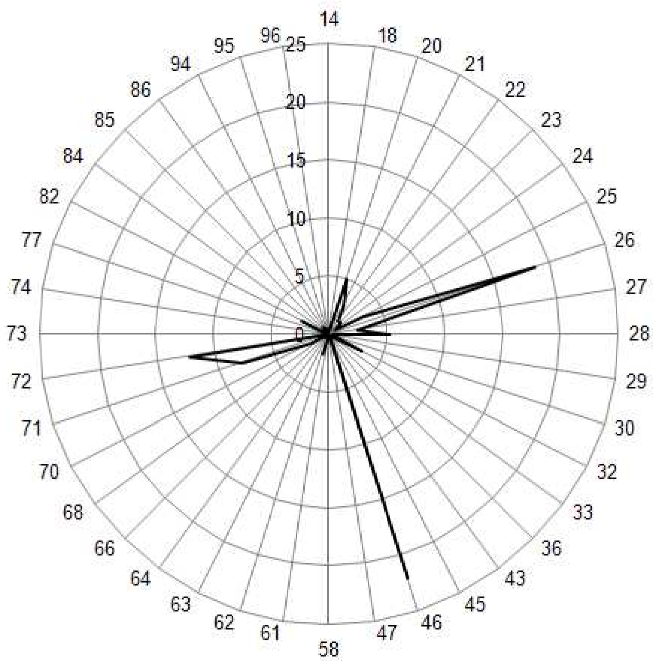
(1) Application field

Diverse sources distinguish nanotechnology companies by *application fields* (see section 5.2.3). Unfortunately, the sources employ different application fields with partially significant overlaps and it is often unclear which definition underlies each field. In this book, this hinders an accurate and consistent assignment of application fields. To avoid imprecision as well as to allow for a common ground, in this section of the book, NACE-codes are employed. NACE-codes give information on the branch(es) a company is active in. Each company can be assigned to one or more branch(es). As each branch is again split into various subcategories, it also occurs that within one NACE-code a company is active in various subcategories.

Fraunhofer, ISI (in: NOYONS ET AL. 2003: 100-101, see section 5.2.2). Patent families are reduced to one entry.
[46] The data is merged by name and location matching.
[47] HOPPENSTEDT usually includes companies with an annual sales volume of at least one million Euro and/ or at least ten employees (see HOPPENSTEDT: Firmenadressen für Direktmarketing, http://www.hoppenstedt-adressen.de/?gclid=CKqQ_5Kur6YCFUYifAodmBQfjg, 10 January 2011). The Hoppenstedt data underlying the doctoral dissertation does include firms with sales and employment data of below these values though.
[48] The data is merged using name matching.
[49] The data is merged using name matching.

NACE-codes are available for all companies considered in this section. Figure 2 (below) depicts their distribution. Note that if a company is assigned to various NACE-codes, all NACE-codes are considered. However, if a company specifies various subcategories within one NACE-code, the respective NACE-code is counted only once. Furthermore, due to the fact that some companies are diversifying companies, it might be that some of the given NACE-codes refer to areas which have no connection to nanotechnology. Thus, the overall picture might be slightly distorted.

Figure 2: Distribution of NACE-codes, in percent.



Source: Own compilations based on dataset 1[50]*.*

In the figure above, each two-digit number refers to one NACE-code (the complete translation into textual language is given in section 5.2.4). In summary, few NACE-codes stand out. Table 4 (below) lists the top ten “chosen” NACE-codes in the sample.

[50] See section 2.3.3.

Table 4: Top ten chosen NACE-codes in the sample, in percent[51].

Rank	NACE-code	Percentage
1	46 Großhandel (ohne Handel mit Kraftfahrzeugen)	22.025%
2	26 Herstellung von Datenverarbeitungsgeräten, elektronischen und optischen Erzeugnissen	18.741%
3	72 Forschung und Entwicklung	12.175%
4	71 Architektur- und Ingenieurbüros; technische, physikalische und chemische Untersuchung	7.934%
5	28 Maschinenbau	5.335%
6	20 Herstellung von chemischen Erzeugnissen	5.062%
7	25 Herstellung von Metallerzeugnissen	3.283%
8	32 Herstellung von sonstigen Waren	3.146%
9	21 Herstellung von pharmazeutischen Erzeugnissen	2.736%
10	82 Erbringung von wirtschaftlichen Dienstleistungen für Unternehmen und Privatpersonen a. n. g.	2.599%

Source: Own compilations based on dataset 1.

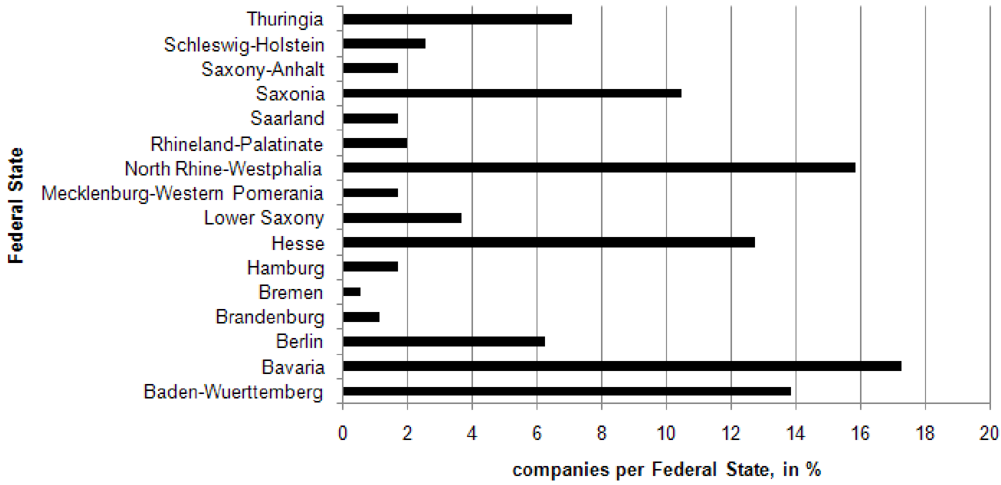
Obviously, wholesale ("46 Großhandel (ohne Handel mit Kraftfahrzeugen)") has the strongest position possibly reflecting the growing (international) importance of nanotechnology which is also suggested by several web pages[52]. Else, it is striking that three of the ten NACE-codes mostly referred to *either* address the chemical *or* the pharmaceutical industry. This again emphasizes the roots of nanotechnology (see p. 4, Figure 1). As also other areas are touched upon (e.g. engineering), the distribution furthermore depicts the diversity of the companies involved in nanotechnology. It seems to confirm the previous statement of nanotechnology affecting a wide range of industries. However, due to the assignment of companies to multiple NACE-codes, and the mentioned, inherited risk of some NACE-code having no direct connection to nanotechnology, all percentages should be treated with care. Some indications might rather be a statistical artifact than a portrayal of reality.

[51] In this section, in line with related works such as CANTNER ET AL. (2005), decimals are rounded to three decimal places. For computations, the exact values are used.
[52] See for example NANO-HANDEL.DE: Nanotechnologie für Handel, Gewerbe & Industrie, http://nano-handel.de/handel-gewerbe-industrie.php?shop=d94d09f05797f3373b00a925db57181c, 10 September 2011.

(2) Location

Next to the application fields, the *location* of the companies is recorded. Figure 3 (below) depicts the distribution of the companies over the 16 Federal States in Germany.

Figure 3: Distribution per Federal State, in percent.



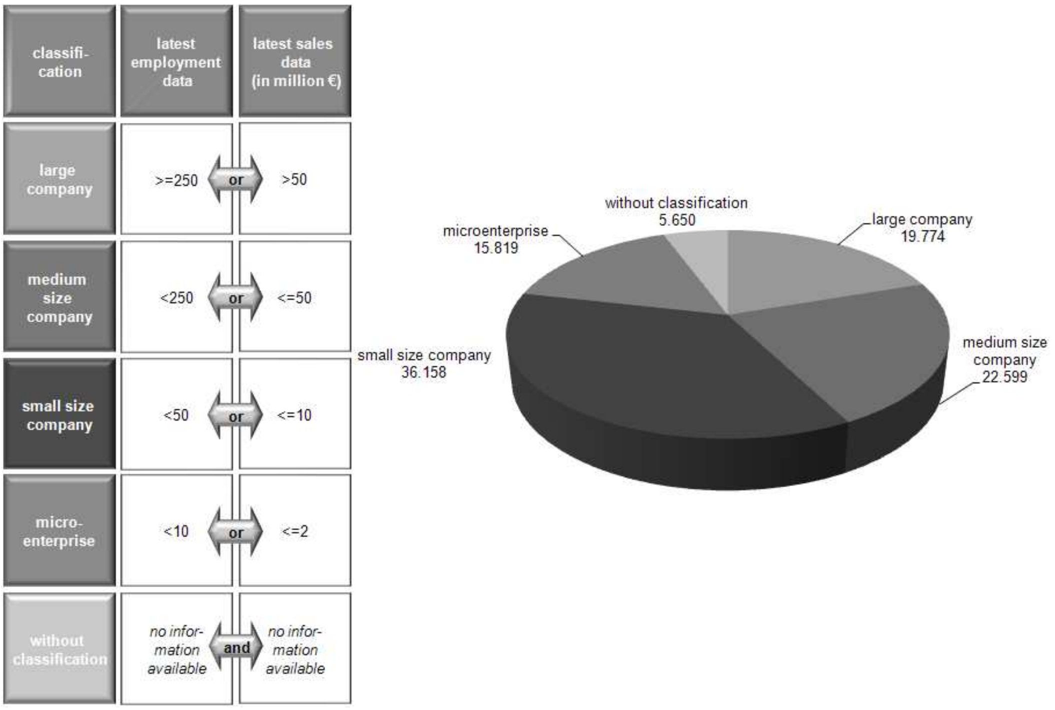
Source: Own compilations based on dataset 1.

Approximately 70% of the companies in the dataset are located in Bavaria, North Rhine-Westphalia, Baden-Wuerttemberg, Hesse and Saxonia. The remaining Federal States are represented in the database to a lesser extent.

(3) Company size

Since companies dealing with nanotechnology can be very diverse in terms of their size, in this paragraph, the distribution of *company sizes* in the sample is presented. Five groups are differentiated: 1. large companies, 2. medium sized companies, 3. small sized companies, 4. microenterprises and 5. companies for which a classification cannot take place due to missing data ("without classification"). The boundary values for the determination of the first four groups are taken from the European Commission 2006 report (see EUROPÄISCHE GEMEINSCHAFTEN 2006: 14). The fifth group is a residual which is added for the purpose of this book. Concerning the assignment to the five groups, slight deviations from the European Commission 2006 report exist (for instance, the report provides an alternative assessment of the company size including a company's balance sheet). In this book, companies are classified according to the annual employment *or* the annual sales data. Which of these indicators is actually taken depends on which data is more recent. If for the same year employment data and sales data indicate different classifications, the "higher" classification is selected. Figure 4 (below) depicts the distribution of company sizes in dataset 1.

Figure 4: Distribution by size following the classification of the EUROPÄISCHE GEMEINSCHAFTEN (2006: 14), in percent.



Source: Own compilations based on dataset 1 and EUROPÄISCHE GEMEINSCHAFTEN (2006: 14).

As shown by Figure 4 above, the 354 companies of dataset 1 divide up into 19.774% large companies, 22.599% medium-sized companies, 36.158% small companies and 15.819% microenterprises. 5.650% of the firms cannot be assigned a size, as for these firms, neither employment nor sales data is available. As only companies with foundry dates after 1978 are considered and large companies often have their roots in the 18th or 19th century, the share of large companies is rather small.

(4) Time frame of analysis

When performing a survival analysis, it is necessary to define a time span in which the analysis takes place. In this work, the *time frame of analysis* is the interval of [1978, 2009]. The choice of the starting date is guided by the (often) perceived “inception” of nanotechnology. The ending date is determined by data availability. Within this time frame, companies “enter” and “exit” the study (see paragraph (5) and paragraph (6) in this section), resulting into different duration times (see paragraph (7) in section 2.3.4 (this section)).

(5) Time of study entry

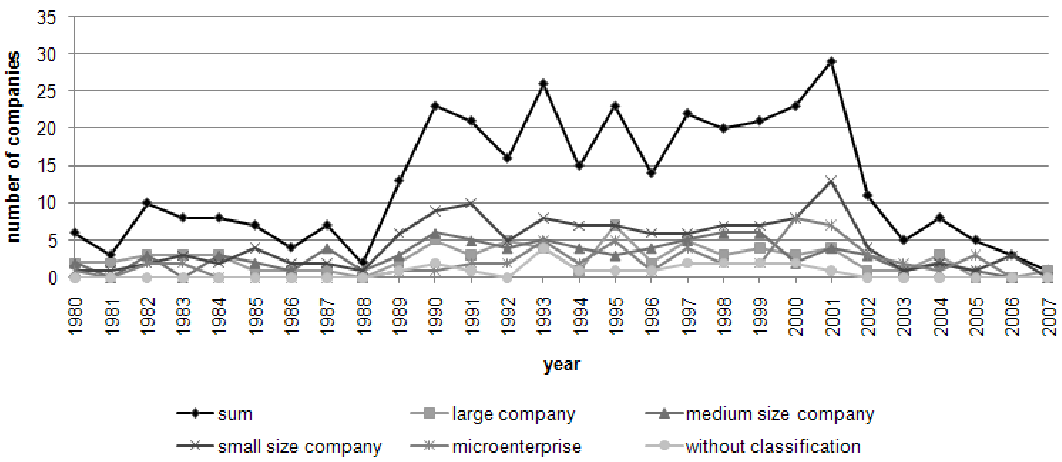
In studies of industry-dynamics, the *time of study entry* often refers to the time of market entry, i.e. to the point of time at which a company begins producing a certain product, such as automobiles. At the time of their market entry, companies are furthermore assumed to either be equipped with pre-entry experience, for example in the sense of being a spin-off, or not.

However, the case of nanotechnology is more complex: Due to the heterogeneity of products it is difficult to capture the exact dates at which companies begin producing their nanoproducts. Sometimes, nanoproducts are not even depicted as such. Furthermore, not all companies in the sample sell and/ or even manufacture products. There are few companies which – for instance – are consultancies. Since for the majority of companies it is *not* possible to extract the point of time at which the specific company begins its work in the area of nanotechnology, in this section of the book, the foundry year ($t_{foundry}$) is used as the time of entry into nanotechnology. The usage of the foundry date comes along with the implication that in the survival analysis, the time a company actually is “alive” is referred to and not the time a company *deals* with nanotechnology. A company may be founded at time $t_{foundry}$ but may actually begin its work in the area of nanotechnology at time ($t_{foundry}+t_1$).[53] It also has to be kept in mind that many companies in the original dataset – especially large companies – have their roots in the 19th or even in the 18th century. While treating the foundry date as the entry time of a company into nanotechnology may be reasonable in those cases in which the foundry date lies between 1978 and 2009, for the companies with earlier foundry dates this approximation is highly disputable: it is not realistic that companies are involved in nanotechnology before 1978. This is why companies with foundry

[53] An analogous pattern is applied for the exit time: a company may still exist at t_2 but may have ceased its work in the area of nanotechnology before.

dates before 1978 are excluded from analysis. Figure 5 (below) illustrates the number of companies per company size and foundry year.

Figure 5: Foundry years.



Source: Own compilations based on dataset 1.

As the course of the sum shows, foundry dates vary strongly in the dataset. There are no foundries before 1980. Until 1988 foundries are rather rare. Afterwards they occur more frequently.[54] Between 1990 and 2001, in total, there are at least 14 foundries per year. After 2001, foundries decline strongly which could demark the end of a first foundry boom. In the years 2007 until 2009 there are no foundries; this is explained by the time of data collection which took place in 2007 and earlier.

At the time of their entry, companies in the dataset are either (or not) equipped with *pre-entry experience*. Paragraph (8) (in this section) delivers information, under which circumstances a company is declared to be a company *with* pre-entry experience and under which circumstances it is labeled to be a company *without* pre-entry experience.

(6) Time of study exit

From the view of 2009, a company may have followed different courses after its foundry. Depending on the pursued course, the *time of study exit* is defined differently. In this book, three “general” categories of courses are distinguished (see Table 5 below).

[54] This seems to match the statement of BOEING (2008: 3) who says that by science, industry and research policy and supported by media since the 1990s, nanotechnology is propagated to be the next great high technology influencing many areas of life.

Table 5: Company statuses in the database, in percent.

Category	Company status (by 2009)	Percentage
1	Active company	77.684%
2	Merger/ acquisition	15.536%
3	Market exit	6.780%

Source: Own compilations based on dataset 1.

Obviously, by the end of 2009 (the end of the study), the majority of companies (77.684%) is still active. 15.536% of the companies fall into the category "mergers and acquisitions". 6.780% of the companies exit the market before 2009. The diverse categories (including *how* in each category the time of study exit is retrieved) are explained in more detail now.

Category 1) Active company

Companies in this category are active until and in the complete year of 2009. This implies that companies exiting the market in 2010 or later (in the sense described by the following category 3) are also declared as active companies in 2009. If a company is still active throughout the complete year of 2009, the time of study exit or the reference year is set to 2009; afterwards the respective companies are no longer under observation.

Category 2) Merger/ acquisition

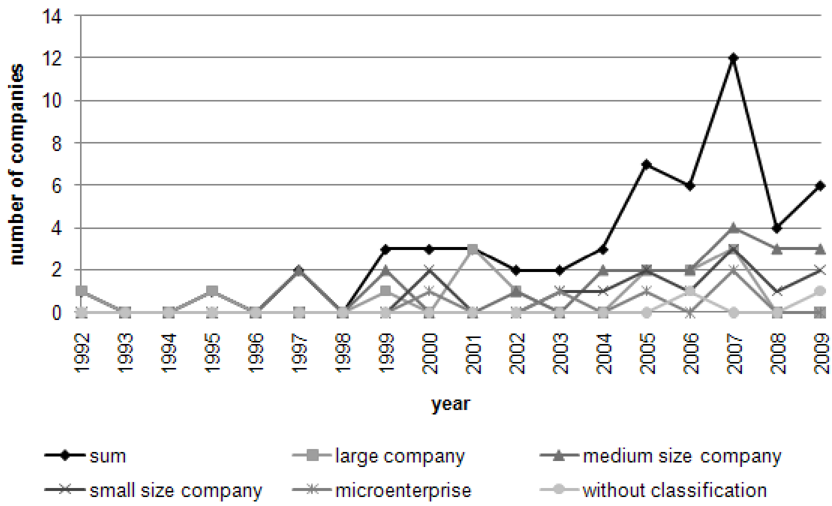
Some companies (1.412%) merge with, some are acquired by other firms (14.124%). Companies in this category are not labeled as "exiting" firms (in the sense described by the following category 3) as they may not categorically be regarded as unsuccessful: Mergers and acquisitions may not necessarily be a sign of failure – they may also be a sign of a firm's attractiveness. In connection with acquisitions, BUENSTORF (2007: 191-193) states:

> "[…] Acquisition may prevent an impending bankruptcy, but alternatively it could be motivated by expected synergies, allowing the original investors to cash in on their investments. Acquisition by a larger competitor is known to be an attractive exit strategy for startups in many high-technology sectors including biotechnology and software. Qualitative evidence also suggests that several promising German laser firms were acquired by competitors. […]"

In the sample, deduced from company information provided by LEXISNEXIS, approximately half of the companies which merge with or are acquired by another company seem to be successful prior to their merger or acquisition. Their merger or acquisition could be due to the fact that they are regarded as specialists, (world) market leaders in a specific segment and/ or leading firms in a particular technology. Even if they are *not* explicitly mentioned to be leading companies, other firms are denoted to be profitable and/ or expanding at least, so they could be regarded as an attractive acquisition candidate. In other cases, companies appear to be innovative or are perceived to be in a good condition, but seem to need another company as an investor. In approximately one fifth of the cases, indications, such as leaving limited partners, losses/ deficits, sales, underpricing of a company or layoffs, *could* point to an unsuccessful company. In a few cases, rather neutral reasons, as for example having a mainstay in Germany or focusing on core competencies, are also mentioned. However, sometimes a company's history prior to the merger and acquisition is not determinable.

To justify the "intermediate character" of mergers and acquisitions, the time of study exit or the reference year is set to the year of merger or acquisition and the respective company is censored this year (in paragraph (7), section 2.3.4 (this section) the reason for censoring is described in detail). Figure 6 (below) shows the course of mergers and acquisitions by company size.

Figure 6: Distribution of mergers and acquisitions per company size.



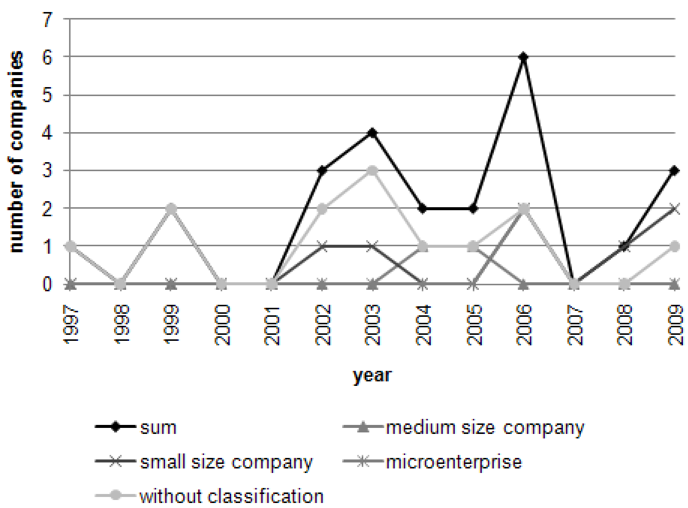
Source: Own compilations based on dataset 1.

As can be deduced from the above Figure 6, the number of mergers and acquisitions per year is initially low, but then increases from 1998 on. Especially in 2007 a peak is reached: in this year, twelve companies either merge with or are acquired by another company.

Category 3) Market exit

Category 1 and category 2 describe cases in which companies do not (really) fail. Companies are either active in the same form as they are originally established or they exist in connection with another company in the sense of a merger or acquisition. However, by 2009, some companies are definitely no longer existent on the market. They actually failed. Three subcategories of such market exits are distinguished: insolvencies (2.542%), liquidations (0.847), and companies which have ceased to exist (3.390%)[55]. In all cases, the time of study exit or the reference year is the year of e.g. insolvency: If – for example – a company declares its insolvency in March 2004, then 2004 is the year of exit, no matter whether the company is active in January or February. Figure 7 (below) depicts the distribution of exits in the sample (distinguishing amongst company size).

Figure 7: Distribution of market exits per company size.



Source: Own compilations based on dataset 1.

As can be deduced from Figure 7, amongst the companies which exit, there is not a single company classified as being "large". It is also obvious that furthermore, there are no exits before 1997. Beginning in 2001, the number of exits increases, mounting up to six exits in 2006. The greater number of exits could be associated with a first – though small – shakeout-phase. However, altogether, there are only 24 exits amongst the 354 companies of dataset 1, so the number is quite small.

[55] These categories are "inherited" from HOPPENSTEDT, which is one data source used to complete the company data in this section.

(7) Duration time and (right-)censoring

The computation of the duration time of companies is straightforward: it is the delta between the time of study exit and the time of study entry. The decisive difference lies in the “treatment” of study exit types: Depending on the kind of study exit, censoring is either done or not.

> “Censoring is defined as when the failure event occurs and the subject is not under observation [...]” (CLEVES ET AL. 2008: 29)

In this book only *right*-censoring occurs meaning that “[...] the subject participates in the study for a time and, thereafter, is no longer observed. [...]” (CLEVES ET AL. 2008: 30). Specifically, right-censoring is applied when a company is still active at the ending point of the study (2009) or if a company experiences a merger or acquisition. It is *not* applied, when a company has actually failed. In other words, right-censoring is used to distinguish companies which have not failed from those which have. Not distinguishing these two categories would mean to declare the ending point of the study and, respectively, the point of merger or acquisition as the point of a company’s market exit. It would therefore imply that active companies and companies which experienced a merger or acquisition experienced a failure event. Table 6 (below) illustrates the computation of duration times in case of right-censored and uncensored data.

Table 6: Computation of duration times, censoring.

Category	Company status 2009	Computation of duration	Right-censoring
1	Active company	$(2009 - t_{foundry})$	Yes (m.s.t.[56]) → 0
2	Merger/ acquisition	$(t_{merger/\ acquisition} - t_{foundry})$	Yes (m.s.t.) → 0
3	Market exit	$(t_{exit} - t_{foundry})$	No (e.s.t.[57]) → 1

Source: Own determination.

The “rules” according to which the companies are assigned to the categories are explained in more detail in the following paragraphs.

Category 1) Active company: computation of duration time and censoring

The duration time on the market is the difference between 2009 and the foundry year ($t_{foundry}$) of the respective companies. In case of active companies, the duration time equals a *minimum survival time* (m.s.t.): Companies in this category are *at least* $(2009 – t_{foundry})$ years active. Since the study ends in 2009, implying that the respective companies are no longer under observation afterwards, the companies are right-censored in 2009.

[56] m.s.t.=minimum survival time.
[57] e.s.t.=exact survival time.

Category 2) Merger/ acquisition: computation of duration time and censoring

In case of mergers and acquisitions, the duration time is the difference between the year of merger or acquisition ($t_{merger/\ acquisition}$) and the foundry year ($t_{foundry}$). Furthermore, the companies are (comparable to the active companies described in category 1) declared as "right-censored"; the duration time again equals a *minimum survival time* (m.s.t.) of the companies on the market.

Category 3) Market exit: computation of duration time and censoring

If a company exits the market due to insolvency, liquidation or because it ceases to exist, the duration time of the company is the year of exit (t_{exit}) minus the foundry year ($t_{foundry}$). Contrary to the first two exit cases described in category 1 and category 2, the company is *not* censored – implying an *exact survival time* (e.s.t.).

(8) Pre-entry experience

As mentioned in paragraph (5), at their time of entry (approximated by the foundry year), companies may either be equipped with pre-entry experience or not. Table 7 (below) shows how many companies are equipped/ are not equipped with pre-entry experience.

Table 7: Types of entry in the database, in percent.

Category	Type of market entry	Percentage
1	companies with pre-entry experience	34.746%
2	companies without detectable pre-entry experience	65.254%

Source: Own compilations based on dataset 1.

For 34.746% pre-entry experience can be determined. For 65.254% of the companies pre-entry experience is *not* detectable. They are considered as "inexperienced". The "rules" according to which the companies are assigned to the two categories are explained now.

Category 1) Companies with pre-entry experience

Pre-entry experience is assumed to be present if the company is a *diversifying* entrant, a *spin-off* and/ or has an *experienced founder*. It is important to note that, for example, a company may have an experienced founder and at the same time be a spin-off company – it may even occur that additionally, the company is considered a diversifying company. This should neither be seen as a weakness nor as a mistake in declaration. Per definition, all three kinds of pre-entry experience are supposed to be uncorrelated in order to cover a large part of possible sources for pre-entry experience. For the analysis, this overlapping effect or rather "multiple counting" is negligible as all three mentioned groups are summarized under the term “companies with pre-entry experience”. The following paragraphs give details on how each subcategory of pre-entry experience is defined.

Subcategory a) Diversifier

Diversifying companies are companies which first produce one product and at a later point of time additionally produce another product. Previous “industry-dynamic” studies therefore identify diversifying companies by analyzing their product portfolio. CANTNER ET AL. (2006: 59) for example record whether firms produce automobiles but origin from other industries. However, as mentioned before, the studies analyze *industries* and not *technologies*. The case of nanotechnology proves to be more complex: Instead of a single, clearly defined product there are many heterogeneous products ranging from surfaces to cosmetics. As there is no unique definition of nanotechnology, it is furthermore difficult to isolate “true” nanoproducts from possibly “false” ones. Sometimes a products’ name might indicate a relation to nanotechnology while at the same time the product has nothing to do with nanotechnology in a proper sense. An example is “Magic Nano Bad- und WC-Versiegeler” and “Magic Nano Glas- und Keramikversiegeler” – cleansing products sold in a German discounter in 2006.[58] According to later examinations there were no nanoparticles in the sprays – the name merely related to a very thin film that the spray left on the surface after application[59]. Furthermore, since some products are still in their developmental phase, assigning a time of market entry is impossible. Even for the products which already are on the market, it is mostly unfeasible to find out about the time of market entry. This information is necessary though in order to find out whether a company producing one nanoproduct and one “none”-nanoproduct is a diversifying entrant into nanotechnology. If the non-nanoproduct is on the market first, then the company is a diversifying entrant. If the opposite is true, the company is *not* a diversifying entrant. In this case, the company would first produce the nanoproduct and later produce the non-nanoproduct. All reasons hinder the product-based identification of diversifying entrants.

Due to the mentioned difficulties, in this work, another approach is used involving patent applications. For each company in the set it is found out, whether the company has applied for a nano patent at the DPMA[60], the EPO or the WIPO[61]. If a company holds a nano patent, in a second step it is investigated whether the company also holds patent applications in areas *other* than nanotechnology. Table 8 (below) lists the cases which can occur.

---

[58] See CHARISIUS, H.: Falsche Furcht, http://www.heise.de/tr/Falsche-Furcht--/blog/artikel/71929, 18 June 2009.
[59] See BUNDESINSTITUT FÜR RISIKOBEWERTUNG: Nanopartikel waren nicht die Ursache für Gesundheitsprobleme durch Versiegelungssprays!, http://www.bfr.bund.de/cd/7839, 18 June 2009.
[60] Deutsches Patent- und Markenamt (engl. German Patent and Trade Mark Office).
[61] World Intellectual Property Organization.

Table 8: Determination of diversifiers.

Category	Patent applications	Diversifier
1	A company does not hold any nano patent applications.	No
2	A company holds nano patent applications, but does not hold any other patent applications "outside" nanotechnology.	No (the company specializes on nanotechnology)
3a	A company holds nano patent applications and holds patent applications "outside" nanotechnology as well. *The priority year of the oldest nano patent application lies further in the past than the priority year of the oldest non nano patent application.*	No
3b	A company holds nano patent applications and holds patent applications "outside" nanotechnology as well. *The priority year of the oldest non nano patent application lies further in the past than the priority year of the oldest nano patent application.*	Yes

Source: Own determination.

Inevitably, detecting diversifying companies using patent applications bears the risk of not finding all patent applications of a company. This is especially true for large companies with more than one identification number under which they apply for patents. Second, the classification of diversifiers by patent applications leads to the unavoidable problem that companies without patent applications can never be classified diversifier.

Subcategory b) Spin-off companies

Spin-offs are new companies emerging from an existing institution, for example a company or a university.[62] In the sample, approximately one third of the spin-offs (31.169%) seem to have their roots primarily in academic R&D (university). Further 29.870% prove to have their roots mostly in non-university R&D, such as in one of the Fraunhofer institutes, Max-Planck institutes or else. 22.078% of the spin-offs seem to mainly derive from companies. Some spin-offs appear to originate from cooperations between two kinds of institutions, for example academic R&D and companies. However, unfortunately, for few spin-offs, the roots are not determinable.

[62] In the dataset, spin-offs are identified using various sources listing spin-offs. Additionally, spin-offs are identified over the internet: The companies' name in connection with the term "spin-off" or "Ausgründung" is searched for. It is assumed that companies provide information on their spin-off roots as usually, they seem to be particularly proud of this kind of origin: Possibly, they aim to attract customers by pointing out their experience and expertise in the field gathered before their actual foundry.

Subcategory c) Experienced founders

Experienced founders are entrepreneurs, who previously lead a company or possibly owned a part of a company. The company of previous employment does not necessarily have to be a nanotechnology company.[63] Unfortunately, for the greatest share of companies, information on the companies' founder is not retrievable.

Category 2) Companies without detectable pre-entry experience

A company is classified as a company "without detectable pre-entry experience" or as "inexperienced" if pre-entry experience as described before is not detectable.

(9) Post-entry experience

In this book, it is assumed that companies entering earlier (or respectively firms which are *founded* earlier) have gathered more post-entry experience compared to firms entering later. Accordingly, diverse entry cohorts are distinguished in the analysis (see Table 9 below).

Table 9: Distribution over entry cohorts.

Cohort	Foundry years	No. of companies
1	[1978, 1992]	128
2	[1993, 1997]	100
3	[1998, 2009]	126

Source: Own compilations based on dataset 1.

The cohorts are chosen such that they are approximately of the same size and that they include a minimum number of events (necessary for performing the survival analysis): To secure reliable results in the regression analysis, the number of exits (for example insolvencies) should not fall below a certain boundary. Otherwise, effects in this cohort are at risk to be biased.

(10) Technological know-how

In the survival analysis, patent applications serve as an indication of technological know-how. Table 10 (below) shows the distribution of technological know-how (patent applications) in dataset 1.

[63] To secure maximum identification, founders are identified in two steps. First, web pages of companies are browsed as partially, these contain information on their founder(s). If this is not the case, in a second step the companies' name together with the term "founded by" or "Gründer" is searched for.

Table 10: Distribution of patent applications, in percent.

Category	Type of market entry	Percentage
1	companies with technological know-how	17.797%
2	companies without detectable technological know-how	82.203%

Source: Own compilations based on dataset 1.

17.797% of the companies hold nano patent applications either at the DPMA or at the EPO/ WIPO. For 82.203% companies, patent applications cannot be determined. Patent families are considered so that multiple counting of the same patent application does not occur.

2.4 Methodological approach

The following remarks serve to briefly sketch the methodological approach, concretely, the survival analysis, applied in section 2.5. In general, a survival analysis determines the influence of defined factors on the life span of individuals. According to CLARK ET AL. (2003: 232) a survival analysis is necessary, when at the end of the follow-up some of the individuals have not had the event of interest and therefore their true time to event is unknown. The authors mention that the second reason for performing a survival analysis is given because survival data is usually not Normally distributed, but is skewed, comprising many early events and few late ones. In this book, the survival analysis consists of two steps: Kaplan-Meier estimates and a (Cox) regression analysis. Both steps are described in the subsequent two subsections.

2.4.1 Kaplan-Meier estimates

Kaplan-Meier estimates allow for a visual representation of the survival probability. CLARK ET AL. (2003: 233) define the survival probability as follows:

> "[…] The survival probability (which is also called the survivor function) S(t) is the probability that an individual survives from the time origin […] to a specified future time t. […]"

Differently put, following CLEVES ET AL. (2008: 7) "[t]he survivor function reports the probability of surviving beyond time t. […] The survivor function is a monotone, nonincreasing function of time". The survival probability per time since entry can be read from the diagram. When two or more survivor functions are drawn into one diagram, for example the survivor function for companies with and for companies without detectable pre-entry experience, according to CANTNER ET AL. (2006: 51) "[…] [t]he validity of the differences of the survivor curves can be further substantiated by a statistical test. […]". In particular, the authors use the log-rank test which – following CLARK ET AL. (2003: 235), "[…] is the most widely used method of comparing two or more survival curves. […]" Concretely, according to

SWEENEY ([NO YEAR PROVIDED]: 13) "[t]he log rank test tests the null hypothesis that […] two groups have the same hazard of failure with a series of k contingency tables (where k is the number of distinct failure times). Stata returns a summary of those contingency tables."[64]

2.4.2 Regression analysis: PH- and AFT-models

When a *simultaneous* treatment of factors is asked for, especially when *predictions* of the hazard[65] or of the survival times are desired, other methods than Kaplan-Meier estimates are indispensable. In this respect, a regression analysis (or, in other words, (semi-)parametric modeling) is a commonly used tool.

When attempting to conduct such kind of "multivariate modeling", it is important for the covariates to meet certain conditions. Next to the implicit assumption of the subjects being representative of a wider population, BRADBURN ET AL. (2003B: 605) mention that data from an adequate number of subjects has to be present: If the number of observations is small this affects the reliability of the estimates and when fitting a model, the impact of the covariates is too imprecise for deriving reliable statements. Furthermore, the authors point out that small datasets may not have sufficient power for detecting a covariate with significant impact on the survival. PEDUZZI ET AL. (1995: 1503) designate a number of 10 EPV[66] as indispensable for retrieving reliable results; saying that if fewer events occur the results of a proportional hazard analysis should be treated with caution as the model may not be valid. Having ensured these conditions are met, the modeling process may take place.

Typically, when performing a survival analysis, the true distribution underlying the data is unknown. "Standard models" are therefore employed which are "adjusted" to the underlying data material, mostly a sample.[67] Table 11 (below) shows the standard models supported by STATA.

---

[64] For further details on the log rank test see CLEVES ET AL. (2008: 123-124).

[65] CLARK ET AL. (2003: 233) describe the hazard as follows: "[…] The hazard is usually denoted by h(t) or λ(t) and is the probability that an individual who is under observation at a time t has an event at that time. Put another way, it represents the instantaneous event rate for an individual who has already survived to time t. Note that, in contrast to the survivor function, which focuses on not having an event, the hazard function focuses on the event occurring. It is of interest because it provides insight into the conditional failure rates and provides a vehicle for specifying a survival model. In summary, the hazard relates to the incident (current) event rate, while survival reflects the cumulative non-occurrence." The hazard depicts the instantaneous rate of failure. The shape of a hazard function may vary over time; it may increase, decrease, remain constant or take serpentine shapes (see CLEVES ET AL. 2008: 7-8).

[66] Events per variable.

[67] In principle, it is possible to create a custom-tailored model. However, finding a mathematical model which reflects the monitored "true" hazard or survival patterns is a task which is usually impossible to achieve. Custom-tailored models are also difficult to verify as it is hard to prove that the correct distribution is assumed.

Table 11: Overview on PH[68]- and AFT[69]-models.

	Semi parametric	Fully parametric
PH Model	Cox	Gompertz Exponential Weibull
AFT Model		Lognormal Log-logistic Generalized gamma Exponential Weibull

Source: Own illustration based on StataCorp LP (2007: 307).

Basically, these models can be distinguished amongst *two* criteria: One criterion concerns the so-called "assumption of proportionality" which Bradburn et al. (2003a: 432) – in connection with the Cox model – explain as follows:

> "[...] The covariates [...] act multiplicatively on the hazard at any point in time [...]: the hazard of the event in any group is a constant multiple of the hazard in any other. This assumption implies that the hazard curves for the groups should be proportional and cannot cross [...]"

PH-models need to fulfill the assumption of proportionality. AFT-models on the other hand include a so-called "acceleration factor" which is computed using the covariates. The acceleration factor has the effect of stretching or shrinking the *survival curve* along the time axis by a constant relative amount (see Bradburn et al. 2003a: 434).

The second criterion concerns the "parameterization" of the baseline hazard[70] (determining whether the model is a semi or a fully parametric model). The difference is depicted by Bradburn et al. (2003a: 432):

> "[…] The key difference between the two [parametric PH models and the Cox (PH) model] is that the hazard is assumed to follow a specific statistical distribution when a fully parametric PH model is fitted to the data, whereas the Cox model enforces no such constraint. […]"

As can be read from the above Table 11, there is no semi parametric AFT model. However, under certain circumstances it would appear useful to have such kind of model available – for example when one single covariate displays nonproportionality while all others in the model fulfill the assumption of proportionality. In this case, modifications of the Cox model can be introduced which take care of the nonproportional covariate.

[68] Proportional hazard.
[69] Accelerated failure time.
[70] See section 2.4.2.2 for more detailed remarks on the baseline hazard.

2.4.2.1 *“Guidelines” for the adequate model choice*

Following previous remarks, choosing the adequate model in principle is guided by two criteria: The first question is how the hazard “behaves” over time (determining whether the model is a PH- or an AFT-model). Testing the assumption of proportionality for its applicability therefore is a key aspect.[71] The other question is whether the baseline hazard is specified (determining whether a semi or a fully parametric approach should be preferred). In general, according to CLEVES ET AL. (2008: 233),

> “[t]he PH metric is used mainly as an analog to Cox regression when the researcher wishes to gain insight into the actual risk process (the hazard function) that causes failure and to gain insight into how the risk changes with the values of covariates in the model. As with Cox regression, little attention is paid to the actual failure times, and predictions of these failures are seldom desired. […]”

On the other hand, the AFT metric is applied, when the aim mainly is to predict failure times or the logarithm of failure times (see CLEVES ET AL. 2008: 233).

When – after testing for the assumption of proportionality – it is clear whether to prefer a PH- or an AFT-model, and if the baseline hazard is parameterized, the question that might remain is which model amongst the PH- or AFT-models explains the observed patterns best. In this respect, BRADBURN ET AL. (2003B: 608) suggest several possibilities: In order to find the most suitable model one – according to the authors “informal” – approach is to plot the (smoothed) empirical hazard or the cumulative hazard against those estimated by the model. Another approach presented by BRADBURN ET AL. (2003B: 608) is to use the so-called AIC[72]. The AIC is a numerical value which is computed the following way (see CLEVES ET AL. 2008: 273):

$$\mathrm{AIC} = -2(\ln \mathrm{L}) + 2(\mathrm{k} + \mathrm{c}),$$

“k“ is the number of model covariates and “c” is the number of model-specific distributional parameters. The distribution with the smallest AIC is the model best fitting the data (see CLEVES ET AL. 2008: 273). In conclusion, concerning the identification of the correct parametric model, BRADBURN ET AL. (2003B: 608) critically remark:

> “[…] In the PH framework, it may be clear that none of the parametric models suggested here or elsewhere adequately capture the distributional form of the data. In such cases, the more flexible Cox model is the obvious choice. Commonly used parametric models in the AFT framework are argu-

[71] STATA yields various options to do so: graphical (“stphplot” and “stcoxkm”) and statistical (“estat phtest”) evaluation tools. For a more detailed description of the three tests see section 5.2.6.
[72] Akaike information criterion.

> ably more flexible than those available in the PH framework, and so fitting a parametric AFT model is another option."

In sum, the presented "guidelines" for choosing the most adequate model are rather "vague" and may only be seen as "hints". They cannot and should not be regarded as fixed rules leading to the perfect solution: A well fitting adjusted standard model may not necessarily be the model truly underlying the basic population. To ease comparability, finally, the decision for or against a model may be influenced by comparable research works applying a distinct model.

2.4.2.2 Brief outline of the Cox model

In order to get a better idea of the models and in anticipation of the later model choice (the model choice is described in detail in section 2.5.2.1), the Cox model is briefly sketched. Mathematically, the hazard function of the Cox model is defined as follows (see CLEVES ET AL. 2008: 152):

$$h(t|x_j) = h_0(t) * e^{(x_j \beta_x)};\ j = 1, \dots, n$$

$h_0(t)$ ... baseline hazard, β_x ... regression coefficients, x_j ... covariates

Basically, the equation on the right hand side consists of two parts: the baseline hazard and the exponential function. In the exponential function, the regression coefficients β_i indicate the "importance" of the covariates x_i. They are estimated in the Cox regression (see STATACORP LP 2007: 126). On the other hand, similar to the intercept in the linear regression, the „baseline hazard" depicts the "basic" hazard for dying when all other covariates are equal to zero (see ZIEGLER ET AL. 2004: 2). In the Cox model, it can be any function of time, merely the effect of the covariates is parameterized which is also why the Cox model is called a semiparametric model (see NOACK 2008: 5).

As mentioned earlier, when a covariate displays nonproportionality, instead of fitting a "regular" Cox model, a *stratified* Cox model may be fitted. In the stratified regression, while the coefficients β_i are constrained to be the same, the baseline hazard is allowed to differ by group (see CLEVES ET AL. 2008: 152):

$$h(t|x_j) = h_{01}(t) * e^{(x_j \beta_x)}, \text{if j is in group 1}$$

$$h(t|x_j) = h_{02}(t) * e^{(x_j \beta_x)}, \text{if j is in group 2}$$

According to BRADBURN ET AL. (2003B: 609) the stratified covariate has to be categorical (or has to be categorized). Furthermore, BRADBURN ET AL. (2003B: 609) say that "[…] more importantly [it] has no estimated effect size provided when forming the strata of a stratified model, and thus is suitable only for covariates that are not of primary interest. […]". Since this extension to the Cox model takes care of nonproportional covariates while also including covariates which fulfill the assumption of proportionality, it can be regarded as being in an intermediate position between the PH- and AFT-approaches (see p. 37, Table 11).

2.4.2.3 Final model assessment

After model application, it is valuable to run several examinations in order to verify whether the model is specified correctly. The first three tests – Cox-Snell, Martingale and Deviance residuals – are based on the computation of residuals. In brief, *Cox-Snell* residuals are useful to assess the overall model fit (see CLEVES ET AL. 2008: 206, STATACORP LP 2007: 170). Is the model suitable, the Cox-Snell residuals should have a standard hazard function equal to 1 and the cumulative hazard should be a straight 45° line (see CLEVES ET AL. 2008: 214). *Martingale residuals* on the other hand are defined by CLEVES ET AL. (2008: 208):

> "[Martingale residuals] can be interpreted simply as the difference between the observed number of failures in the data and the number of failures predicted by the model."

They can be computed to determine the functional form of the covariates which are included in the model (see CLEVES ET AL. 2008: 206). It is pointed out that due to their range sometimes Martingale residuals are difficult to interpret and for this reason, deviance residuals are preferred (see STATACORP LP 2007: 172, 339). For the fully parametric models the procedure of retrieving the correct functional of the covariates resembles the procedure as described in the context of the Cox model. The difference is that Martingale-*like* residuals are computed. STATACORP LP (2007: 339) explains:

> "We use the term "martingale-like" because, although these residuals do not arise naturally from martingale theory for parametric survival models as they do for the Cox proportional hazards model, they do share similar form."

According to CLEVES ET AL. (2008: 217) "[i]n evaluating the adequacy of the fit model, it is important to determine if any one or any group of observations has a disproportionate influence on the estimated parameters." There are various options to check for these so-called "outliers". *Deviance residuals* for example may serve to examine model accuracy and to identify outliers (see CLEVES ET AL. 2008: 206). They are a rescaling of the Martingale (or Martingale-like) residuals; they are symmetric about zero and stronger resemble the residuals from linear regression (see STATACORP LP 2007: 173, 339).

Furthermore, CLEVES ET AL. (2008: 197-198) recommend the link test: Having fitted a model, the test verifies that the coefficient on the squared linear predictor is insignificant. In other words this implies:

> "[...] **linktest** is based on the idea that if a regression is properly specified, one should not be able to find any additional independent variables that are significant except by chance. linktest creates two new variables, the variable

> of prediction, _hat, and the variable of squared prediction, _hatsq. The model is then refit using these two variables as predictors. _hat should be significant since it is the predicted value. On the other hand, _hatsq shouldn't, because if our model is specified correctly, the squared predictions should not have much explanatory power. That is we wouldn't expect _hatsq to be a significant predictor if our model is specified correctly. So we will be looking at the p-value for _hatsq. […]"[73]

At the same time, CLEVES ET AL. (2008: 198) emphasize that the link "[…] test is weak in terms of detecting the presence of omitted variables. […]". The link test may therefore only reasonably indicate model misspecification if it is certain that all essential covariates are included in the model. This constraint should be kept in mind when interpreting the test results. The test may not be considered a suitable quality indicator for the "general" model assessment if there is reason to believe that essential covariates are missing in the model.

2.5 In-depth analysis: significance of experience and knowledge

In this work, the analyzed objects are nanotechnology companies. Their individual life span or duration time is determined as described in Table 6 (see p. 30). It is examined, how three factors (pre-entry experience, post-entry experience and technological know-how) possibly influence the survival of a company or its hazard of leaving the market. Table 12 (below) summarizes the factors or rather *covariates* under investigation and presents the testing concept used.

Table 12: Examined covariates possibly influencing the survival of companies.

Covariate	Testing concept
pre-entry experience	companies *with* pre-entry experience vs. companies *without* detectable pre-entry experience
post-entry experience	early entry vs. late entry
technological know-how	companies *with* technological know-how vs. companies *without* detectable technological know-how

Source: Own determination[74].

Analogously to CLARK ET AL. (2003), not all firms experience a market exit (see p. 27, Table 5): The majority of companies are still active at the ending point of the study (2009). This makes a survival analysis attractive.

[73] UCLA ACADEMIC TECHNOLOGY SERVICES: Stata Web Books Regression with Stata Chapter 2 – Regression Diagnostics, http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm, 27 July 2009.
[74] The work of CANTNER ET AL. (2006) served as a source of inspiration.

2.5.1 Nonparametric analysis: Kaplan-Meier estimates

In this work, survival curves are computed in general and per covariate (see Table 12 above). For the "computation" of Kaplan-Meier estimates, the time frame of analysis has to be synchronized and split into periods. The length of these periods is determined by the duration time of the companies (see p. 30, Table 6). With the duration time and the corresponding information about a company's status available, the data is prepared for further analysis. For the 354 companies of dataset 1 first descriptive statistics show that in total, there are 24 failures implying that 24 of 354 companies exit the market according to category 3 (see p. 30, Table 6). In general, the observation begins at the synchronized time t_{eoe}=0 (=time of earliest observed entry) and the longest observed duration time is t_{loe}=29 years (=time of last observed exit (non-censored or censored)). More detailed statistics on dataset 1 can be deduced from Table 13 (below).

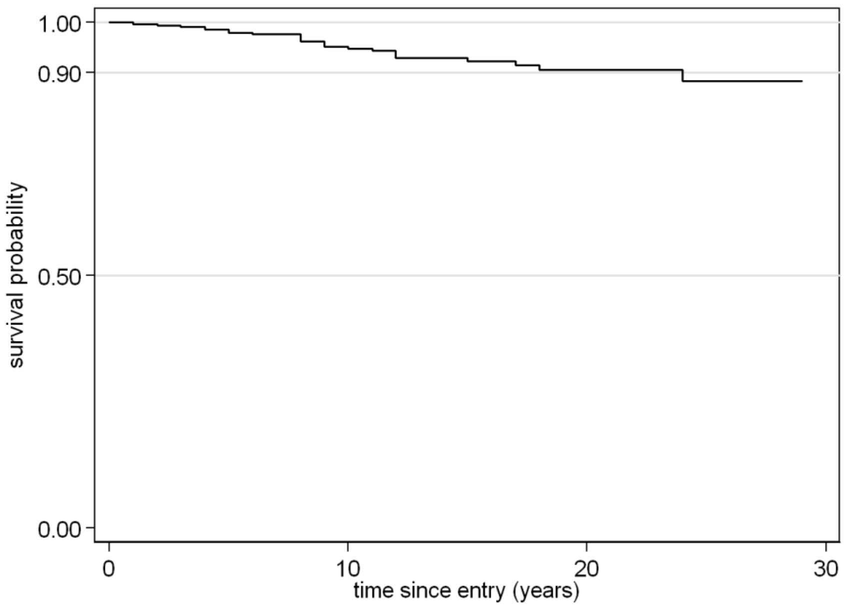
Table 13: Further statistics of the dataset.

		per subject			
	total	**mean**	**min**	**median**	**max**
# subjects	354				
# records	354	1	1	1	1
(first) entry time		0	0	0	0
(final) exit time		13.644	1	13	29
subjects with gap	0				
time at risk	4830	13.644	1	13	29
failures	24	0.068	0	0	1

Source: Own compilations based on dataset 1.

Of the 354 observations or subjects, due to the previously described synchronization of the time axis, all companies have an entry time of t=0. As a consequence, the mean, minimum, median and maximum entry times are zero. The minimum (final) exit time is 1, the maximum (final) exit time is 29 years leading to a mean (final) exit or duration time of 13.644 years. The median (final) exit time is 13 years. There are no gaps in the dataset – if companies exit the study, they do not re-enter. As in the dataset, the 24 failures (or market exits) are marked with "1" while the censored data (active companies, mergers or acquisitions) are marked with "0" (see p. 30, Table 6), the minimum value is "0", the maximum value per subject is "1" and the mean (which in case of a binary variable is of limited meaning) is 0.068. In most of the cases, the data is censored, so the median is "0". In order to get a first visual impression on the general survival pattern, Figure 8 (below) depicts the Kaplan-Meier curves (*not* distinguishing amongst covariates).

Figure 8: Kaplan-Meier survival curve.



Source: Own compilations based on dataset 1.

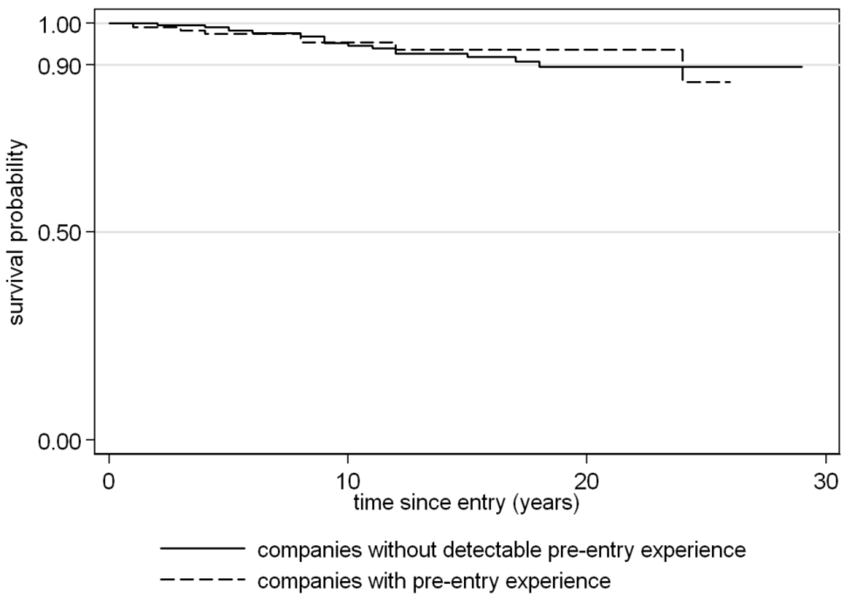
For each cut in the timeline, the number of exits and censors is counted so that on this basis, the cumulative survival can be computed. As can be deduced from Figure 8, the survival probability declines; the distribution is right-skewed. As most of the cases are censored, the survival curve never reaches a survival probability of 0.000% and rather remains at a high level of around 90.000%. This hinders the computation of a mean survival time as the computed mean is always underestimated. After artificially "extending" the survival function to zero or, in other words, extrapolating the survival function until it reaches zero (which is necessary in order to compute a mean survival time which is realistic at least to some degree), the mean survival time[75] for the 354 companies of dataset 1 is computed to 233.005 years. Of course, this value is also fictive – but it indicates that the survival time of companies (at least from the perspective of 2009) is rather long. However, the above statistics as well as the graph do not yield any detailed information on the effect of the covariates "pre-entry experience", "post-entry experience" and "technological know-how" on the survival of companies. Therefore, Kaplan-Meier curves are drawn in the following and statistics are given separately for each of the covariates depicted in Table 12 (see p. 41).

[75] See section 5.2.5 for further information on the computation of mean survival times in dataset 1.

(1) Pre-entry experience

Kaplan-Meier curves are drawn for companies *with* pre-entry experience and for companies *without* detectable pre-entry experience. Figure 9 (below) shows the Kaplan-Meier curves for both.

Figure 9: Kaplan-Meier curves for pre-entry experience.



Source: Own compilations based on dataset 1.

The Kaplan-Meier curves above show that companies with pre-entry experience and companies without detectable pre-entry experience have similar survival probabilities in the first (approximately) 12 years of their existence. Afterwards, having pre-entry experience seems to “pay off”; at least until age 24 the survival probability for companies with pre-entry experience is slightly higher. It then drops below the survival probability of companies without detectable pre-entry experience. The log rank test supports the impression of qualitatively similar curves: With a P-value of 0.866 the null hypothesis of the two curves being the same cannot be rejected. The extended means picturing the long term perspective are depicted by Table 14 (below).

Table 14: Pre-entry experience, extended mean.

	# Subjects	Extended mean
companies with pre-entry experience	123	172.175
companies without detectable pre-entry experience	231	262.906
total	354	233.005

Source: Own compilations based on dataset 1.

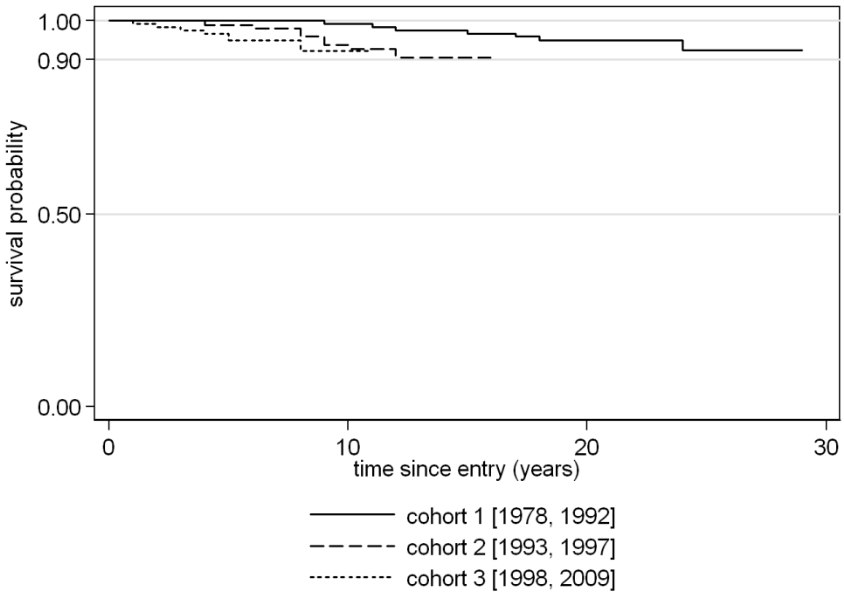
Table 14 shows that in the *long run*, companies with pre-entry experience have a lower survival probability expressed in terms of a shorter expected lifespan. Based on the overall view, hypothesis 1 *can neither be confirmed nor rejected:* The survival probability of companies with pre-entry experience and companies without detectable pre-entry experience is comparable in the first (approximately) 12 years. Afterwards the survival probability of companies with pre-entry experience is higher than in case of the companies without detectable pre-entry experience. At age 24, the survival probability of companies without detectable pre-entry experience is higher.

Recalling the findings of comparable studies as for example CANTNER ET AL. (2006: 51-52), on first sight, the findings may seem astounding. The results based on dataset 1 indicate that – while being of lesser relevance in the short-term perspective – having an experienced founder, being a spin-off or a diversifier is especially relevant in the mid-term perspective. It seems as if – for example at the time of their inception – decisions have been made in companies with pre-entry experience which secure their mid-term survival. Companies without detectable pre-entry experience on the other hand seem to have partially lacked "having put things on the right track"; their survival probability is a little lower in the mid-term view. However, *if* companies without detectable pre-entry experience manage to survive (and as can be seen the probability is high), they seem to be more successful in the long run.

(2) Post-entry experience

As mentioned before (see section 2.3.4, paragraph (9)), the data is split into three cohorts according to the foundry year of the companies. Kaplan-Meier curves are drawn for each cohort. Figure 10 (below) shows the respective Kaplan-Meier curves.

Figure 10: Kaplan-Meier curves for post-entry experience.



Source: Own compilations based on dataset 1.

In terms of post-entry experience, the Kaplan-Meier curves clearly show that companies entering earlier have a higher survival probability. With a P-value of 0.020 (implying that the null hypothesis of the groups having the same hazard can be rejected) the log rank test also shows that the differences between the curves are not only noticeable visually but are also statistically significant. The extended means depicting the long term perspective confirm this picture (see Table 15 below).

Table 15: Post-entry experience, extended mean.

	# Subjects	Extended mean
cohort 1	128	370.432
cohort 2	100	160.456
cohort 3	126	135.586
total	354	233.005

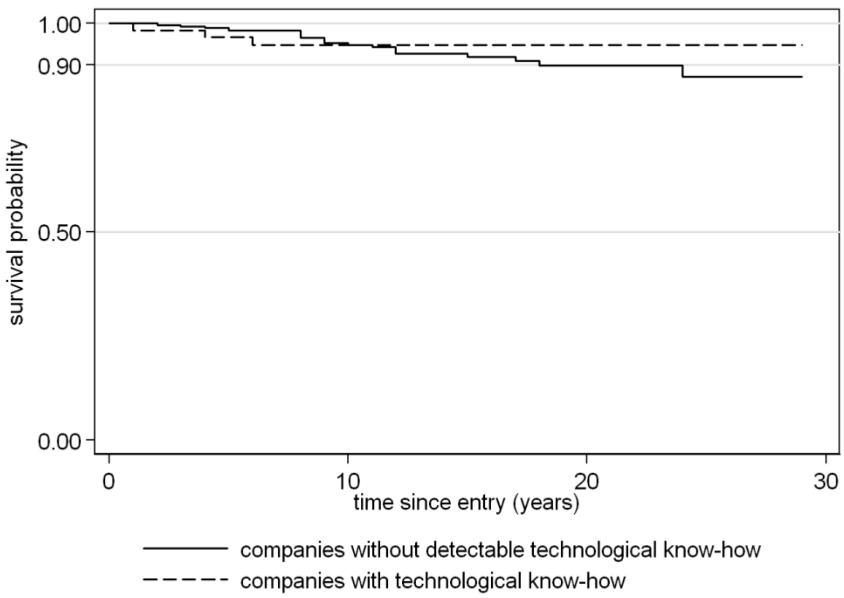
Source: Own compilations based on dataset 1.

Companies in cohort 1 have a higher mean survival time than companies in cohort 2 and companies in cohort 2 have a higher mean survival time than companies in cohort 3. Thus, hypothesis 2 is confirmed. Furthermore, the results are in line with comparable studies in this area, for example CANTNER ET AL. (2006: 52).

(3) Technological know-how

Kaplan-Meier curves are drawn for companies which have technological know-how and for companies for which technological know-how is not determinable. Figure 11 (below) shows the respective Kaplan-Meier curves.

Figure 11: Kaplan-Meier curves for technological know-how.



Source: Own compilations based on dataset 1.

In terms of technological know-how, the Kaplan-Meier curves show that companies *with* technological know-how have a slightly *lower* survival probability in the first couple of years of their existence, but, in the long run, they are more successful in the sense of having a higher survival probability. While – with a P-value of 0.611 – the log rank test outlines that the null hypothesis of the two curves being the same cannot be rejected, the extended means (depicting the long term perspective) support the relevance of technological know-how in the long term perspective (see Table 16 below).

Table 16: Technological know-how, extended mean.

	# Subjects	Extended mean
companies with technological know-how	63	551.705
companies without detectable technological know-how	291	211.322
total	354	233.005

Source: Own compilations based on dataset 1.

Altogether, hypothesis 3 can therefore be considered *partially confirmed*: In the long run, nanotechnology companies have a higher survival probability if they are equipped with technological know-how. Else, they have a lower survival probability.

What could be hypothesized is that – at the beginning – companies without technological know-how mainly focus on their business enrolment in general, not in particular on following up on patent applications. This could lead to a quick business start-up which possibly makes the respective companies more successful in the short-term perspective. To the contrary, companies with technological know-how may put more energy into patenting i.e. into developing a unique technology. This could lead to the failure of some companies as the technological evolution might head into another direction or as possibly, costs imposed by the patenting procedure might be too high to bear. However, if companies with technological know-how manage to survive (and the survival rate is above 90%) they seem to have gained a solid fundament to build upon in the future.

2.5.2 Semiparametric analysis: Cox regression

Kaplan-Meier estimates give a first impression regarding the influence the factors pre-entry experience, post-entry experience and technological know-how exert on the survival of companies. According to the findings presented above, *in the long run*, the presence of pre-entry experience leads to a lower survival probability, while earlier entry and the presence of technological know-how leads to a higher survival probability.

While these findings may offer a first valuable insight, all three covariates are treated separately when assessing their relevance for the survival of companies. It cannot be determined whether the presence of pre-entry experience for example influences the survival of companies to a greater extent than the presence of technological know-how or whether the effect of one (or both) factors may be compensated when a company enters early. Therefore, in the following, a regression model is applied to *simultaneously* estimate the impact of the three covariates (see p. 41, Table 12) on the survival (time) of companies or rather, the *hazard* for companies of leaving the market.

2.5.2.1 Legitimation for the application of the Cox model

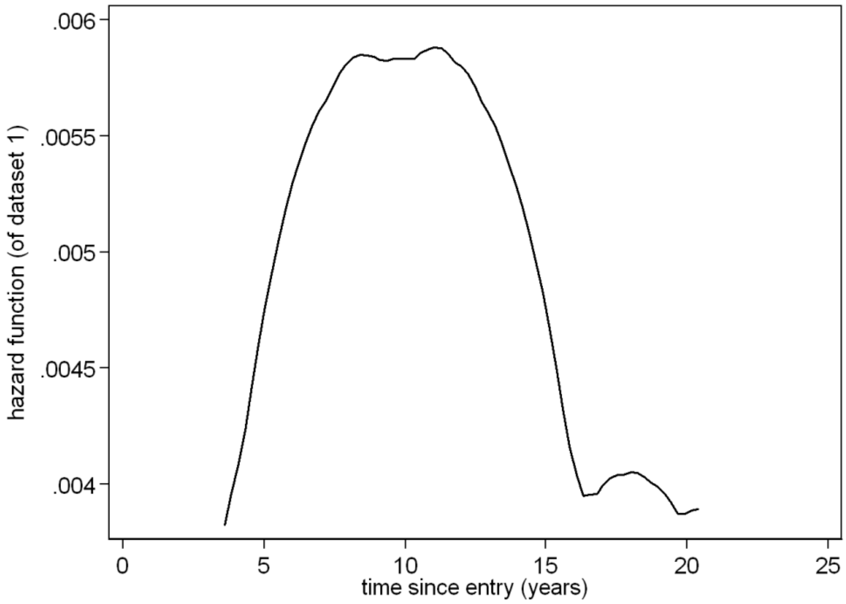
Before being able to perform a regression analysis, it has to be assessed which regression model is best to apply: It has to be evaluated, whether the assumption of proportionality is reasonable to assume. Second, it needs to be clear whether the baseline hazard is specified.

To secure objectivity, in this book, the assumption of proportionality is tested with a non-graphical method, the so-called "estat phtest".[76] According to the global test, there is no evidence that the assumption of proportionality is violated, so altogether, findings point to the application of the PH-metric. A slight violation can be deduced from the covariate technological know-how (Prob>chi2 is significant in this case). Recalling the Kaplan-Meier estimates for the covariate "technological know-how" there is evidence that being equipped with technological know-how is advantageous in the long run. At the inception, not being equipped with technological know-how seems rather favorable.

The other question to answer is whether a semi- or a fully parametric approach should be preferred. In the semi-parametric approach, the baseline hazard is not parameterized. To the contrary, in fully parametric approaches the baseline hazard is specified. In order to find out which distribution amongst the fully parametric models best describes the observed hazard – recalling section 2.4.2.1 – according to BRADBURN (2003B: 608) one option is to compute the AIC. Following the AIC, in the PH-metric, the Weibull model would be the most suitable option. Due to the fullfillment of the assumption of proportionality, in general, the AFT-metric appears to be less suitable. However, if an AFT-model was recommendable to use, the AIC would suggest the Lognormal model (see section 5.2.7.2). Further following BRADBURN ET AL. (2003B: 608), another possibility for finding the most suitable model is to draw the smoothed hazard based on dataset 1 (see Figure 12 below) and to hold it against the hazard estimated by several models (see section 5.2.7.2).

[76] The detailed results of the estat phtest test are shown in section 5.2.7.1.

Figure 12: Smoothed hazard.



*Source: Own compilations based on dataset 1***.**

As Figure 12 shows, in comparison to the hazard rates of the fitted models (see section (2)), the smoothed hazard underlying the sample data does not suggest the hazard in the sample to underlie any particular distribution. Initially, the hazard in the sample seems to rise and appears to reach a first local maximum around age 8. The hazard then decreases a little and remains approximately constant around age 10, but then increases again to reach a global maximum around age 12. Afterwards, the hazard declines strongly. A second “local” maximum is reached around age 18. This course is not depicted by any of the fitted models, not even for the Weibull or the Lognormal model just previously recommended by the AIC. However, even if the *sample* suggested a particular distribution to be present, this distribution might not agree with the (mostly unknown) distribution of the *basic population*: It is possible for the hazard in the sample to strongly deviate from the hazard in the basic population. Resultingly, applying a fully parametric model to the sample data bears the risk of applying a model which does *not* reflect the true hazard. This will unavoidably lead to wrong conclusions in the attempt of making predictions.

Altogether, *because* findings suggest the application of the PH-metric to be reasonable and *because* within the PH-metric (and also within the AFT-metric) findings do not indicate a particular hazard to be present, the decision is to apply a (stratified) Cox model to the data. This decision seems furthermore substantiated because in section 2, the focus is on the actual risk process; the aim is to measure the effect of selected covariates on the exit hazard of firms. The prediction of failure times – which according to CLEVES ET AL. (2008: 233) would suggest the application of the AFT-metric – is not in the center of attention. Another reason for applying the Cox model is that it is extremely flexible: it allows for the consideration of potential nonproportional covariates. Lastly, the Cox model is frequently used in comparable studies, so comparability between studies is eased.

2.5.2.2 Findings of the Cox regression

In this work, three covariates x_1 (pre-entry experience), x_2 (post-entry experience) and x_3 (technological know-how) are to be evaluated in terms of the influence they (simultaneously) exert on the hazard of firms of leaving the market. To do so, a formal regression analysis (consisting of several covariate constellations) needs to be performed which yields information on the covariates' regression coefficients (β_i) and their significance levels. Since the former test of the assumption of proportionality revealed the PH-approach to be applicable in principle, but also showed that the variable technological know-how displays a slight nonproportionality, in addition to the "original" Cox regression, a stratified Cox regression is applied whenever the covariate technological know-how is considered in a model. In this modified approach the baseline hazard of companies with technological know-how and the baseline hazard of companies without detectable technological know-how are allowed to distinguish from another. Furthermore, the efron method is applied to account for ties occurring in the database.

To estimate the effect of the covariates, diverse Cox regression models are fitted to the underlying data. The models have in common that they are all estimated by partial likelihood using the program STATA 10. Furthermore, each model is based on the set of 354 companies of dataset 1 of which 24 companies "fail" due to insolvency, liquidation or because they cease to exist. The essential difference between the models lies in the covariates they include. Generally, the set of explanatory variables which is considered consists of a dummy variable to indicate the presence of pre-entry experience and three dummy variables for the three entry cohorts. Furthermore, models can either be stratified by the covariate technological know-how (also a binary covariate) or not. In principle, this leads to the combinations listed by Table 17 below.

Table 17: Model configuration.

	Model					
	M1	**M2**	**M3**	**M4**	**M5**	**M6**
pre-entry experience	1	0	1	1	0	1
post-entry experience	0	1	1	0	1	1
stratification (by technological know-how)	no	no	no	yes	yes	yes

Source: Own determination.

Model 1, model 2 and model 3 estimate whether pre- or post-entry experience alone or in combination may account for the observed hazard curve in the sample. In all three models, technological knowhow is not included as a stratification variable: Models 1, 2 and 3 depict “original” Cox regressions. In turn, model 4, model 5 and model 6 include the same variable constellation as model 1, model 2 and model 3, but *additionally* include technological know-how as the stratification variable: Model 4, 5 and 6 depict stratified Cox regressions. Table 18 and Table 19 (below) display the results of the fitted Cox models.

Table 18: Cox regression (efron method for ties).[77]

	Model 1	**Model 2**	**Model 3**
Log likelihood	-130.799	-126.688	-126.502
n	354	354	354
EPV[78]	24/1=24	24/2=12	24/3=8
Prob>chi2	0.865	0.016	0.035
pre-entry experience	0.927 [0.866]	*not included*	0.761 [0.549]
post-entry experience			
(cohort 1)	*not included*	*(1.000)*	*(1.000)*
cohort 2	*not included*	*3.372 [0.045]*	*3.460 [0.041]*
cohort 3	*not included*	*5.470 [0.010]*	*5.778 [0.009]*
technological know-how	not included	not included	not included

Source: Own compilations based on dataset 1.

Table 19: Stratified Cox regression (efron method for ties).[79]

	Model 4	**Model 5**	**Model 6**
Log likelihood	-121.803	-117.541	-117.494
n	354	354	354
EPV[80]	24/1=24	24/2=12	24/3=8
Prob>chi2	0.887	0.014	0.035
pre-entry experience	1.074 [0.886]	*not included*	0.858 [0.761]
post-entry experience			
(cohort 1)	*not included*	*(1.000)*	*(1.000)*
cohort 2	*not included*	*3.519 [0.038]*	*3.551 [0.037]*
cohort 3	*not included*	*5.643 [0.009]*	*5.823 [0.009]*
technological know-how	stratified	stratified	stratified

Source: Own compilations based on dataset 1.

According to the computations, model 1 and model 4 are insignificant as a whole: Prob>chi2 of model 1 is 0.865 and respectively, Prob>chi2 of model 4 is 0.887. Due to their insignificance, these models are of secondary interest and are not explored any further. Regardless of whether technological know-how is considered as a stratification variable or not, pre-entry experience alone may not ex-

---

77 Note that P-values are reported in parentheses [] besides the hazard rate.
78 According to PEDUZZI ET AL. (1995: 1503) a number of 10 EPV is indispensable for retrieving reliable results; saying that if fewer events occur the results of a proportional hazard analysis should be treated with caution as the model may not be valid.
79 Note that P-values are reported in parentheses [] besides the hazard rate.
80 According to PEDUZZI ET AL. (1995: 1503) a number of 10 EPV is indispensable for retrieving reliable results; saying that if fewer events occur the results of a proportional hazard analysis should be treated with caution as the model may not be valid.

plain the observed hazard patterns. These results reflect the findings of the log-rank test performed on the covariate pre-entry experience (see section 2.5.1, paragraph (1)).

The remaining four models all are statistically significant with their Prob>chi2 ranging between 0.014 and 0.035. Amongst these, model 2 only includes the cohort dummies. Following KLEPPER (2002B: 48) cohort dummies are chosen to be equal to unity in case a company is part of the respective cohort and zero otherwise. To avoid multicollinearity between the cohort dummies, one cohort always needs to be omitted in the regression analysis. In the models depicted in Table 18 and Table 19, cohort 1 is chosen to be omitted, so the reported hazard estimates are expressed as a proportion of the hazard of cohort 1. Concretely, model 2 displays that the later the entry, the higher the exit hazard: The results show that companies in cohort 2 face 3.372 times the hazard of companies in cohort 1 and companies in cohort 3 face 5.470 times the hazard of companies in cohort 1. The estimates show that both cohort dummies are statistically significant. This strengthens the impression that earlier entry is immediately connected to gathering more experience which is useful in terms of surviving on the market. The picture seems even a little sharper when additionally considering technological know-how as a stratification variable (see model 5). With comparable hazard estimates, the significance level of the overall model as well as of the cohort dummies is even a little higher. In general, the relevance of the cohort dummies reflects the results of the log rank test on the covariate post-entry experience (see section 2.5.1, paragraph (2)).

Following CANTNER ET AL. (2006: 50FF), post-entry experience is related to the experience companies gather during their operation on the market. In other words, the longer they are alive, the more post-entry experience they collect. Implicitly, this implies that at birth, all firms are equipped with zero post-entry experience. However, at their inception, companies may, for example, have obtained prior knowledge from parent firms or may be equipped with knowledge from other technological fields. In this notion, model 3 expands model 2 by also considering pre-entry experience as a secondary source for obtaining survival-relevant knowledge. Pre-entry experience is a dummy variable which is equal to unity in case of firms with pre-entry experience and zero otherwise. However, as in model 1, pre-entry experience is statistically not significant. In turn, the cohort dummies are statistically significant; their reported hazard rates display a similar picture as in model 2. Being part of cohort 2 implies facing 3.460 times the hazard of cohort 1 and being part of cohort 3 implies facing 5.778 the hazard of cohort 1. Model 6, which also accounts for differences in technological know-how, draws an even sharper picture. The significance levels of the cohort dummies partially are a little higher than in case of model 3. However, model 3 and model 6 have an EPV-ratio of 8 which is rather low, so it might occur that in these models, the results are slightly biased.

To assess possible compensational effects, in two further models the dummy for pre-entry experience appears interacted with the dummy variables for the entry cohorts. One model is stratified by the variable technological know-how, the other model is not stratified. Table 20 (below) displays the results.

Table 20: Interaction effects (Cox regression, efron method for ties).[81]

	Model 7	**Model 8**
Log likelihood	-130.370	-121.275
n	354	354
EPV[82]	24/2=12	24/2=12
Prob>chi2	0.642	0.584
cohort2* pre-entry experience	1.448 [0.559]	1.772 [0.388]
cohort 3* pre-entry experience	1.762 [0.385]	1.744 [0.407]
technological know-how	not included	stratified

Source: Own compilations based on dataset 1.

As can be read from Table 20, neither the models nor any of the interacted cohort dummies display statistical significance. Compensational effects therefore do not seem relevant.

In conclusion, it can be said that, apparently, the presence of *pre-entry experience* is not relevant in terms of explaining the hazard patterns in the underlying sample. Even *if* the P-value displayed significance, the hazard rate would still be close to 1, so the difference in hazard rates between companies with pre-entry experience and companies without detectable pre-entry experience would be marginal. Against the background of the previously drawn Kaplan-Meier curves the results are not surprising: The Kaplan-Meier curves for pre-entry experience display that companies with and companies without detectable pre-entry experience have survival curves which are pretty much alike. Companies with pre-entry experience are better off only for a limited span of time. Overall, the findings of the Kaplan-Meier estimates and the Cox regression lead to the conclusion that hypothesis 1 can neither be confirmed nor rejected by the data.

While pre-entry experience does not seem to be of relevance, the findings indicate *post-entry experience* to be significant: Based on the Kaplan-Meier curves and the regression analysis, earlier entry seems to secure survival or, differently put, reduce the exit hazard of firms. This could be a consequence of older com-

[81] Note that P-values are reported in parentheses [] besides the hazard rate.
[82] According to PEDUZZI ET AL. (1995: 1503) a number of 10 EPV is indispensable for retrieving reliable results; saying that if fewer events occur the results of a proportional hazard analysis should be treated with caution as the model may not be valid.

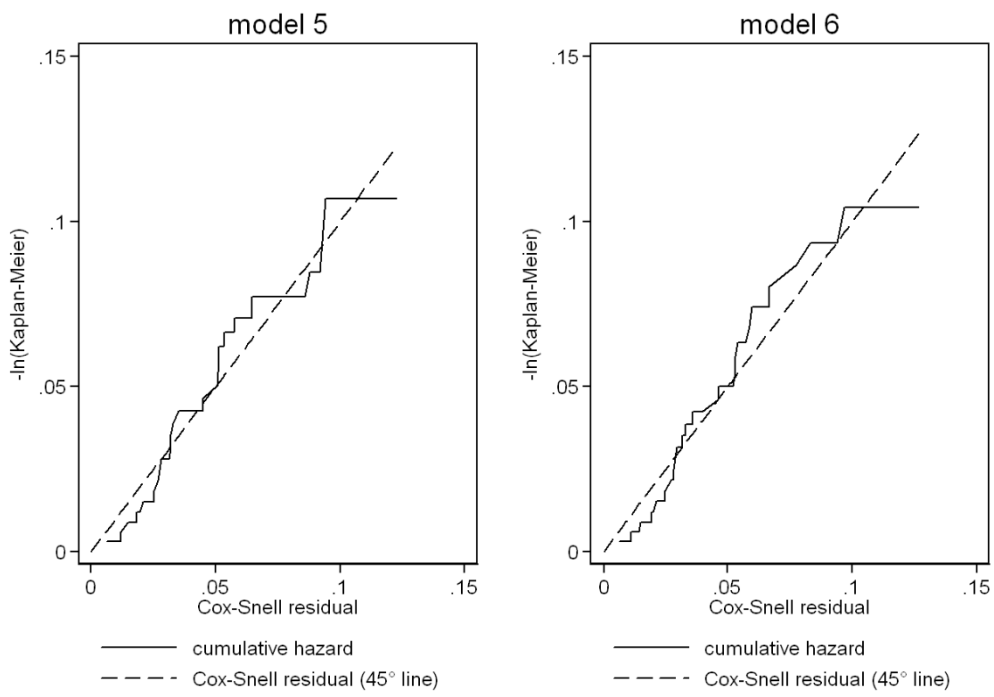
panies already being established on the market; they might for example have secured themselves a fixed client base and know about their customer's tastes. These and similar factors could help them wisely choose the direction to head into. In turn, having former experience in other market segments does not really help increasing the survival probability: nanotechnology seems to be a field of its own – with its own challenges to manage. Altogether, hypothesis 2 is confirmed.

Finally, the covariate *technological know-how* cannot be evaluated with the stratified Cox model because the model is stratified by this variable. While the Kaplan-Meier estimates partially confirm hypothesis 3, based on the regression results, hypothesis 3 can neither be confirmed nor denied.

2.5.2.3 Assessing model adequacy of the Cox model

According to section 2.5.2.2 four of the six presented models display statistical significance. Concretely, these are model 2, model 3, model 5 and model 6. The latter two distinguish from model 2 and, respectively, model 3 in terms of considering the stratification variable technological know-how. It turns out that the corresponding results are even a little sharper. Therefore, models 5 and 6 are investigated in this section. Both models are subject to various tests to assess their precision. At first, Cox-Snell residuals are computed for the assessment of the overall model fit (see Figure 13 below).

Figure 13: Cox-Snell residuals of model 5 and model 6.

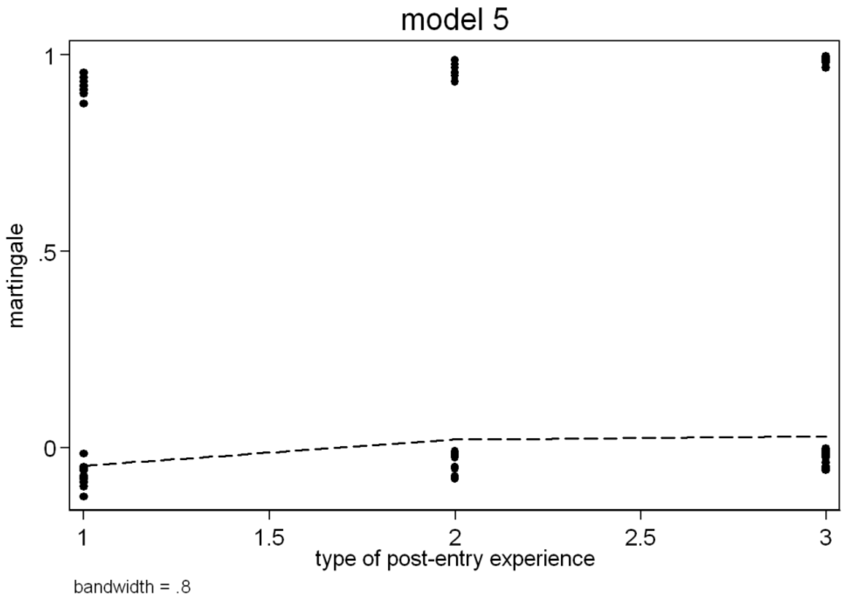


Source: Own compilations based on dataset 1.

According to Figure 13, both models display a pretty well fit. There are only slight deviations from the 45° line.

To further enhance model fit, it can be tested whether including functions of the covariates would be a better choice. Since including a functional form of the binary covariate pre-entry experience seems of limited sense, in the following, only the covariate post-entry experience is examined. In this work, the procedure of how to determine the correct functional form of the covariate is oriented towards the procedure as suggested by CLEVES ET AL. (2008: 209): First, the Cox model is fitted excluding all covariates and the Martingale residuals are computed. In the next step, each covariate is plotted separately against the Martingale residuals. A smoothing is included to ease finding the correct functional form. Figure 14 (below) shows the results. Since the results for model 6 are alike, only the results of model 5 are displayed.

Figure 14: Post-entry experience, Martingale residuals for model 5.

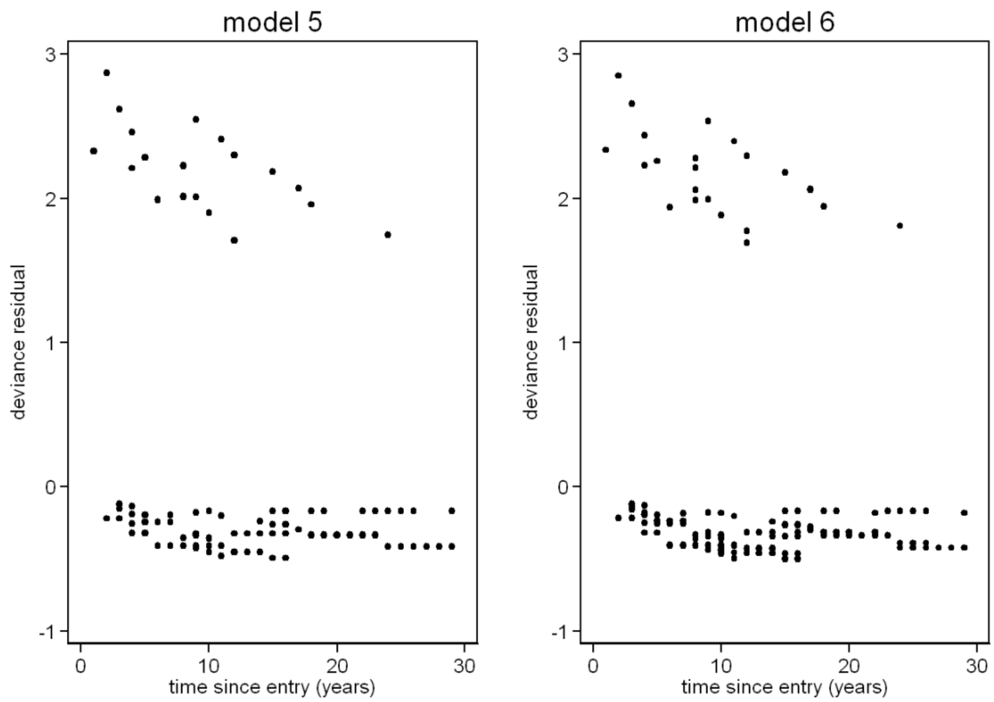


Source: Own compilations based on dataset 1.

If approximately, the smooth function displays a linear form, the transformation is adequate (see STATACORP LP 2007: 172). As can be deduced from Figure 14, in case of post-entry experience, the smooth is roughly linear. A transformation of the covariates is therefore not necessary.

As according to CLEVES ET AL. (2008: 206) and STATACORP LP (2007: 170) deviance residuals may serve to examine model accuracy and to identify outliers, finally, deviance residuals are computed for the models (see Figure 15 below).

Figure 15: Deviance residuals for model 5 and model 6.



Source: Own compilations based on dataset 1.

In both models, there are some outliers in the upper part of the figures. The models do not seem to fit perfectly. Finally, the results of the link test are shown below (see Table 21 below).

Table 21: Link test for model 5 and model 6.[83]

	Model 5	Model 6
Log likelihood	-126.688	-126.538
n	354	354
Prob>chi2	0.016	0.014
_hat	0.923 [0.491]	1.046 [0.387]
_hatsq	0.034 [0.966]	-0.033 [0.963]

Source: Own compilations based on dataset 1.

[83] Note that P-values are reported in parentheses [] besides the hazard rate.

With a P-value of 0.966 in case of model 5 and a P-value of 0.963 in case of model 2, the variable "_hatsq" is not significant. Under the assumption that all essential covariates are included in the model, the model seems to be specified correctly. However, "_hat" is not significant in either case which indicates that the model is not perfect.

In conclusion, based on the results of the evaluation methods, the validity of model 5 and model 6 is given in principal – the overall model fit is pretty well and mathematical transformations of the covariates are not needed. However, outliers do have a certain influence on the results.

2.6 Summary and conclusion

The aim of this book is to explain the evolution of companies in the knowledge-driven area of nanotechnology. In this section three factors – pre-entry experience, post-entry experience and technological know-how – are examined to evaluate their influence on the actual firm survival or, differently put, on the hazard of firms of exiting the market. Concretely, three hypotheses are under investigation.

> ***HYPOTHESIS 1:*** Nanotechnology companies have a higher survival probability if they are equipped with pre-entry experience.
>
> ***HYPOTHESIS 2:*** Nanotechnology companies have a higher survival probability if they are equipped with more post-entry experience.
>
> ***HYPOTHESIS 3:*** Nanotechnology companies have a higher survival probability if they are equipped with technological know-how.

In preparation of the empirical analysis, data from five different sources listing nanotechnology companies is assembled and prepared using HOPPENSTEDT and PATSTAT (VERSION 10/ 2007). A very extensive manual research involving sources as LEXISNEXIS and GENIOS, but also but also a very high number of – for example – corporate web pages, press releases etc. is furthermore conducted to close existent gaps in the database. Altogether, 354 companies (founded between 1978 and 2009) – referred to as "dataset 1" – are processed.

To test the above mentioned hypotheses on the basis of dataset 1, the analysis is divided into two parts: In a first step, Kaplan-Meier estimates are performed. The estimates are used to gain first insights into possible reasons of firm survival. As a simultaneous treatment of covariates is not possible with Kaplan-Meier estimates, in a second step, more substantiated multivariate modeling is done: To find out more about the dependency between pre-entry experience, post-entry experience, technological know-how and firm survival, a Cox regression is performed considering several covariate constellations. Since the covariate technological know-how displays a slight nonproportionality, a *stratified* Cox regression is applied whenever the covariate technological know-how is included in a model. Altogether, six models are considered differing according to the covariates they comprise. Four of the six fitted models display statistical significance: Concretely, these are two “unstratified” models and their “stratified counterparts”. Since on covariate level as well as in terms of the overall model significance the stratified models display slightly improved results, the stratified models (model 5 and model 6) are subject to further evaluation strategies. Altogether, the results of the Kaplan-Meier estimates and all Cox models are summarized in the following. The results are grouped by and organized in the order of the hypothesis.

Specifically, hypothesis 1 moves the factor *pre-entry experience* to the center of attention. In this respect, previous works on industries agree that it does have a profound influence on firms – its presence leading to a lower hazard ratio or a higher survival probability (see KLEPPER 2002A: 661FF, CANTNER ET AL. 2006: 52, 57FF, THOMPSON 2003: 15, 27). The findings for nanotechnology differ from these findings: From the Kaplan-Meier estimates it can be learned that companies *with* pre-entry experience have a higher survival probability in a *limited time span*. Furthermore, the Cox models including the covariate pre-entry experience show that the significance level of this covariate is always rather low. Therefore, it is questionable whether pre-entry experience may explain the survival of companies to a satisfying extent. Altogether, hypothesis 1 can neither be confirmed nor rejected.

In terms of hypothesis 2, which is concerned with the factor *post-entry experience*, previous studies (amongst others) detect that – by the time of the shakeout – earlier entrants have lower hazard rates persisting many years thereafter; on the other hand, hazard rates seem to decline with age in some industries while in other industries age does not exert an influence on the hazard (see KLEPPER AND SIMONS 1999: 36FF). Examining the U. S. automobile, beer brewing and tire industry, HORVATH ET AL. (2000: 18) for instance report elevated hazard rates for cohorts entering late in an industry’s life cycle. Overall, the Kaplan-Meier estimates for nanotechnology companies suggest that companies entering earlier to the technology’s life cycle have a higher survival probability, supporting the findings of previous researchers. “Older” companies are more resistant to failures than newcomers. The Cox models (in which the covariate post-entry experience is included) confirm these findings: the *later* the entry the *higher* the hazard of exiting. In all models, the significance levels of the covariate post-entry experience are very high. Hypothesis 2 is therefore confirmed.

With respect to *technological know-how* acquired through innovative activities (which is in the center of attention in hypothesis 3) previous studies suggest a correlation between innovative activities and the survival of companies. Innovative activities may even compensate for the lack of post- and pre-entry experience (see CANTNER ET AL. 2005: 1). In the area of nanotechnology, a connection between innovative activities and survival is also determinable: The Kaplan-Meier curves show that companies with technological know-how have a higher survival probability in the long run, but a lower survival probability in the beginning. Due to the application of the stratified Cox model, a co-evaluation of the covariate technological know-how cannot take place. Nevertheless, whenever the covariate technological know-how is considered as a stratification variable, the overall model results reach slightly higher significance levels, so it appears as if the covariate at least exerts a little influence. Mainly based on the findings of the Kaplan-Meier estimates, hypothesis 3 can be regarded *partially confirmed.*

Beyond the scope of the previously mentioned models, compensational effects are also subject of investigation. Concretely, in two additional models, the covariate pre-entry experience is interacted with the covariate post-entry experience. However, regardless of whether or not technological know-how is considered as a stratification variable, compensational effects turn out not to be of significance.

In order be able to derive a statement concerning the validity of the models, model 5 and model 6 (the two stratified models with the highest significance levels) are then examined: Cox-Snell residuals are computed to assess the overall model fit. The residuals depict only slight deviations from the 45° degree line – indicating the overall model fit is pretty well. When investigating the Martingale residuals it is revealed that a mathematical transformation of the covariate post-entry experience does not lead to any noteworthy improvement. The covariate may be included in its original form. Finally, the deviance residuals give information on possible outliers and model accuracy. As there are some outliers it can be deduced that the model is not too accurate. This result is confirmed by the link test, which is included last.

While for the time being, the model(s) seem(s) reasonable to consider to explain actual firm survival, for future works in this area it may be reasonable to put further effort into the enhancement of the database: Especially, because more precise data is not available yet, the present survival analysis is based on the time a company actually is “alive” and not on the time a company is actually occupied with nanotechnology. This might distort the results to a certain degree. Furthermore, the EPV-ratio is rather low in all models. The suggested remedy therefore is to continuously track the development of the companies in order to observe long or at least longer term trends. This might lead to an even better explanatory power of the whole model.

Beyond these rather “basic” concerns it should be considered to expand the database to the extent that – next to experience and knowledge – for example size effects or regional- and industry-specific effects are included. For reasons of comparability, originally, at least size effects were meant to be included in the analysis. Unfortunately though, the share of exiting companies proved to be disproportionally high amongst the companies without classification (see section 5.2.8). Excluding these companies from the entire analysis would therefore have led to a dramatic reduction in exiting companies. This again would have impaired the overall explanatory power of the models. Concerning regional effects, Figure 3 (see p. 23) displays that some Federal States such as Bavaria, North Rhine-Westphalia and Baden-Wuerttemberg are strongly represented in the database (at least amongst the 354 companies of dataset 1) while others are not. The dissimilar distribution might be a peculiarity of the sample, but might also indicate that in some Federal States, the infrastructure is more fruitful for companies than in other Federal States. Else, industry-specific factors could play a role in the sense of influencing some companies in the dataset while not influencing others. Figure 2 (see p. 21) hints that – amongst other – the chemical industry is occasionally referred to in dataset 1. It is imaginable that companies which are active in this area are influenced by maybe new guidelines established for the chemical industry. In other words, if an industry suffers a shock or experiences a boom this might exert an influence on the (survival of) companies in the industry no matter if they are involved in nanotechnology or not. All, size-, regional- as well as industry-specific aspects could be subject to further examination exceeding the mere investigation of experience and knowledge for the survival of nanotechnology companies. Having considered further covariates, the approach suggested by BRADBURN ET AL. (2003B: 606) who propose verifying the choice of covariates by a degree of hands-on modeling (that is to add or remove terms in a logical order) should be thought about (the authors do not recommend the adding or removal process to be merely based on statistical significance).

Independent of the suggested extensions, a *survival analysis* always establishes a connection between diverse, possibly very different kinds of factors and survival. The strength of a survival analysis is to be seen in the fact that it yields information of *whether* and *how* the selected *covariates* influence survival, in case of section 2, *actual* firm survival. However, aside from analyzing *actual* firm survival, specifically in the area of nanotechnology, another approach to explain the evolution of companies is valuable to apply simultaneously. This approach specifically addresses the fact that nanotechnology is a *technology*. In the following section 3 this additional approach is presented in detail.

3 IMPORTANCE OF TECHNOLOGICAL FIELDS

3.1 Motivation and aim

The survival analysis performed in section 2 delivers valuable information on *whether* and *how* the factors pre-entry experience, post-entry experience and technological know-how relate to firm survival. *Survival* is thereby defined as *actual* firm survival. Specifically in terms of nanotechnology, another, yet *alternative* perception of survival, relating to the *technological* evolution of firms, is recommendable though. *Why* it is rewarding to focus on the technological evolution of firms and finally, *how* survival can be defined in this respect, is explained in the following paragraphs.

Amongst other, one finding of section 2 is that the corporate landscape underlies remarkable dynamics. Foundries of companies involved in nanotechnology vary strongly over the past years (see p. 26, Figure 5). At the same time, some companies merge with or are acquired by other companies (see p. 28, Figure 6); some even experience a market exit (see p. 29, Figure 7). Furthermore, it is obvious that the application fields of companies, approximated by NACE-codes, differ strongly from another (see p. 21, Figure 2). Under consideration of the changing corporate landscape and the differences in their scope, it seems straightforward to assume that also the *technological fields* touched upon by the nanotechnology companies or – in other words – the *technological orientation* of nanotechnology companies changes over time.

In connection with a potential technological evolution, companies might play different roles. Some companies may face strong competition because they act in the same technological fields as others, whereas other firms may be technologically isolated working in less common technological fields. Similarly, some companies may rather push the development in one or several technological fields while others do not.

Hence, *if* such roles are present, the question is, whether these can account for the *technological* survival of firms within the area of nanotechnology: For example, companies focusing on a selection of technologies which are commonly used by other firms may underlie tough competition, but they might also be less at risk to (technologically) exit as – for instance in case of the unsuccessful pursuit of one technology – theoretically, there are technological alternatives to follow up to or to expand. On the other hand, being specialized in one technology within nanotechnology which no other firm follows up to *could* be helpful to claim an outstanding, potentially leading market position. At the same time, it may be risky if the technological development of nanotechnology is heading into another direction, i.e. is in favor of other technological fields. To the extreme case, a company may be forced to technologically exit the area of nanotechnology.

When it comes down to a closer investigation of the technological evolution, profound information on companies and their technological scope is needed. The information content thereby needs to go beyond stating *whether* (as in section 2) a company is equipped with technological know-how or not. Concretely, information on *which* technological fields a company touches upon needs to be present. Second, an adequate tool for e.g. investigating roles amongst companies in relation to their survival needs to be considered.

To justify these requirements, in section 3, another approach than performed in section 2 is reasonable to pursue, involving the thorough examination of patent applications: A social network analysis[84] is carried out to examine whether the general "choice" of technological fields of a company (deducible from its patent specifications[85]) may account for a company's *technological survival*. Concretely, the underlying question in section 3 is whether a company with *many* technological overlaps to other companies (implying that it possibly underlies stronger competition) is *less* endangered to experience a technological exit from the area of nanotechnology than a company with a *smaller* technological overlap to other companies.

The consideration of patent applications allows for a very *dynamic* definition of entry, exit and survival: Next to the technological fields which are touched upon, patent specifications include – amongst others – the priority date of the patent application. While the foundry date of a company does not need to imply that a company begins its work in the area of nanotechnology, the priority date is a more precise indicator for a company having started its work in this area. By splitting the time span of analysis into diverse periods of time, entry, exit and survival may be defined as follows: A company *enters* into cohort (t+1) *technologically*, if it has not applied for a (nano-)patent in t, but applies for a (nano-)patent in (t+1). Respectively, a company *exits* from cohort t *technologically* if it holds a patent application in cohort t but does not apply for a patent in (t+1).[86] Companies which apply for patents in succeeding cohorts obtain the status "permanent actor", i.e. they *survive technologically*. In comparison to section 2, this approach tolerates *repeated* entries and exits and furthermore allows for a better approximation of the time a company is actually occupied with nanotechnology.

[84] The term "*social* network analysis" is a common term in network analysis. It mostly refers to *willingly* established connections between two or more actors. In line with previous works, the term is adopted, though, in *this* book, connections between actors are of rather theoretical nature.
[85] See section 5.3.1 for an exemplary patent application at the DPMA.
[86] *Technological* exit should therefore not be equated with market exit: A company may technologically exit from the area of nanotechnology but may still be active in areas outside nanotechnology.

3.2 Literature overview and choice of hypotheses

Works applying methods of social network analysis are manifold. This makes it difficult to draw up a shortlist fully reflecting the state of the art. With respect to the analysis performed in section 3.5, therefore, only a selection of studies is presented in this section. Without raising an issue of completeness, the goal is to give an impression of the existing studies and to thereby prepare and explain the choice of hypotheses.

In general, in several works, networks are assigned an essential role for innovation and competition. Alluding to the importance of cooperation within networks, TEECE (1992) for example specifically addresses the relation of competition, cooperation and innovation. He argues that – particularly in fragmented industries – complex forms of cooperation are usually necessary to promote competition (see TEECE 1992: 3). The author states that "[…] advanced technological systems do not and cannot get created in splendid isolation. […]" (TEECE 1992: 22). POWELL ET AL. (1996) on the other hand specifically concentrate on networks of learning in the area of biotechnology. They argue that "[…] when the knowledge base of an industry is both complex and expanding and the sources of expertise are widely dispersed, the locus of innovation will be found in networks of learning, rather than in individual firms. […]" (POWELL ET AL. 1996: 116).

Based on the assumption of networks playing a crucial role for innovation and competition, CANTNER AND GRAF (2006) perform a social network analysis to describe the evolution of the innovator network in the city of Jena, Germany in the period from 1995 to 2001. Constructed by the help of patent data (patent applications for instance contain information on the priority year, the inventors and applicant(s), but also on the technological orientation of patents), they define and analyze two types of networks: The network of technological overlap and the network of personal relations. In both networks, the actors are institutions (called innovators by CANTNER AND GRAF (2006)). In case of the network of technological overlap, ties between institutions stand for a present technological overlap[87] between them. In case of the networks of personal relations, ties between institutions stand for a personal relation[88] between them. Both types of networks are computed for two cohorts: 1995-1997 and 1999-2001. In a first step, they are concerned with structural attributes of the networks. Amongst other, they find out that the networks experience an increase in size from 139 actors in the first to 189 actors in the second cohort, which equals an increase of roughly 36%. Also, in

[87] "**Technological overlap**: Linkages between innovators are formed whenever they patent in the same technological class. This network can be interpreted as the potential for cooperation." (CANTNER AND GRAF 2006: 466).

[88] "**Personal relations** distinguished into:
Cooperation: When there is more than one innovator on a patent, there are as many linkages between all co-applying innovators as there are inventors.
Scientist mobility: Whenever a specific inventor is mentioned on patents applied for by distinct, not cooperating innovators a link between those innovators exists, since the inventor has worked for both." (CANTNER AND GRAF 2006: 466).

both kinds of networks, an increase in the average number of ties between the actors is perceivable.

Next to the previously described attributes, specifically, issues of actors playing different roles within networks are often addressed. Specifically, such roles concern the position of actors within the network and their linkages to other network members. STANLEY (2006) for example takes a closer look at the formation of network patterns and analyzes what types of patterns are likely to emerge given different circumstances. The author suggests taking a closer look at game theory, where the Iterated Prisoner's Dilemma might be useful to explain the binary decision of an actor to either cooperate or to not cooperate in distinct periods of time. Depending on the previous behavior of a potential partner, there might be incentives for an actor to either cooperate or not.

AUTANT-BERNARD ET AL. (2007) study collaboration networks in micro- and nanotechnologies on the basis of R&D projects submitted to the 6th Framework Program. After characterizing the structure of the R&D collaborations between firms with the help of a social network analysis, the determinants of this structure are examined by analyzing the choices of cooperations. The authors find out that network effects are present so that the probability of collaboration is influenced by each individual's position within the network. Furthermore, they discover that social distance is more important than geographical distance and that firms with similar research potential are more likely to work together (see AUTANT-BERNARD ET AL. 2007: 1).

CANTNER AND GRAF (2006) are also concerned with the identification of roles of actors. For Jena, based on the network of technological overlap, CANTNER AND GRAF (2006: 467-469) find that larger innovators form the center of the networks. These organizations prove to have a high technological overlap to each other. Other firms are in the periphery of the networks. Else, over the cohorts, CANTNER AND GRAF (2006) report an increase in network centralization. Altogether they interpret the increase in cohesion as a stronger focus on core competencies and identify the central actors to become increasingly important for the entire network. Concerning the networks of personal relations, CANTNER AND GRAF (2006: 469-470) summarize that large, core actors seem to increasingly focus on formal cooperation while smaller actors in the periphery seem to focus on informal, personal relations.

Specifically with regard to explaining evolution, structural attributes of networks such as entry and exit from cohorts and, respectively, survival can be related to the roles of actors. For the network of technological overlap, CANTNER AND GRAF (2006: 471-472) find that between permanent actors and entrants (which they define as local firm foundings or as firms which relocate), there are more technological connections than between permanent and exiting actors. Furthermore, permanent actors experience a strong growth in their technological overlap over the cohorts. This leads the authors to conclude that permanent actors and entrants tend to increasingly concentrate on the technological core competencies of the network as a whole. For the network of personal relations, CANTNER AND GRAF (2006: 473-474) find that entrants cooperate significantly more with permanent actors than the exiting actors did. They also appear to have more linkages through scientist mobility though this is not to a significant degree.

If networks have such a high impact on innovation and competition, they should be thoroughly observed and analyzed over time – specifically in new and promising technologies like nanotechnology. In this book, therefore, a network theoretic approach is chosen which – due to its well fitting strategic procedure – is inspired by CANTNER AND GRAF (2006): Patent applications are used to construct networks of technological overlap for several cohorts.

Specifically, in this book, *nanotechnology* patent applications are used to construct networks of technological overlap for five subsequent cohorts (altogether comprising the time between 1978 and 2005). Concretely, in each network of technological overlap, the actors are domestic companies applying for a nanotechnology patent in the respective cohort. The ties between the actors stand for a present technological overlap between the companies. In other words (following CANTNER AND GRAF 2006: 466): Whenever two companies apply for a patent in the same technological class, a link between these two companies is drawn. This implies that in each network of technological overlap, all companies applying for a patent in the same technological field are connected to each other. Altogether, 35 technological fields are distinguished following the WIPO IPC-Technology Concordance Table (see SCHMOCH 2008: 9-10 or section 5.3.4).

It is important to note that *because* ties between companies result from a present technological overlap, the connections between companies are of *potential* and (at least mostly) not of *actual* nature (e.g. in the sense of a cooperation). To the utmost, the constructed networks may therefore be regarded as an *indicator* for explaining possibly occurring cooperations. In their work, in addition to the analysis of the networks of technological overlap, CANTNER AND GRAF (2006) included aspects of cooperation by considering networks of personal relations. However, in this book, networks of personal relations are difficult to construct and to evaluate, because – contrary to CANTNER AND GRAF (2006) – the analysis is based on entire Germany instead of one singular town. This particularly raises the issue of correct name assignments: Since the same name may refer to different persons, on large scale it is hard to verify whether two patent applications with the same inventor name actually relate to one or two researchers. Instead of focusing on aspects of cooperation, in line with the overall aim of this book, this section focuses on explaining the technological survival of companies in the area of nanotechnology by using the networks of technological overlap. Concretely, three hypotheses are investigated:

> ***HYPOTHESIS 4:*** The structure of the network of technological overlap changes strongly over the observed cohorts in terms of the actors

which are part of the network and in terms of the intensity of ties between the actors.

HYPOTHESIS 5: In each period of time, in the networks, there are few diversified actors which – by means of social network analysis – are clearly identifiable as core actors with a high technological overlap to other actors. Consequently, there are also companies in the periphery of the network with a small technological overlap to other companies.

HYPOTHESIS 6: The majority of companies in the core of the network in cohort t remain actors of the network in cohort (t+1). The majority of companies in the periphery of the network of cohort t exit after cohort t, i.e. are not part of the network in cohort (t+1).

As mentioned, the hypotheses are tested by means of a social network analysis. Similar to the survival analysis (see section 2), performing a social network analysis requires the availability of an extensive dataset. In the following subsections (see sections 3.3.1 to 3.3.4) the data is described and descriptive statistics are given.

3.3 Database

3.3.1 Data ascertainment

The list below presents a brief overview on the data which is necessary for constructing the network of technological overlap. A detailed description of each of the variables follows in section 3.3.4.

(1) Application authority: In this book, the EPO and WIPO are considered as application authorities. Patents applied for at these organizations are recognized as technologically and economically highly valuable (see FRIETSCH ET AL. 2008).

(2) Applicants sequence number/ Inventor sequence number: These numbers are necessary to identify the first, second etc. applicant/ inventor of a patent application. In this book, only data is analyzed for which the applicants sequence number is greater than zero.

(3) Person country code: In line with the aim of this book, in this work, only patent applications from Germany are considered.

(4) Person name: The person name is either the name of the applicant or of the inventor. In this book, diverse categories of person names are distinguished ("company", "academic R&D", "non-university R&D", "charity", "person" or "other"). The focus is on companies.

(5) Priority date: Patent applications with a priority date between 1978 and 2005 are examined. In preparation of the social network analysis, this period is divided into five cohorts.

(6) Application ID: The application ID allows for the unique identification of a patent. In this book, 1284 patent applications are analyzed (deriving from the companies).

(7) Technological fields: According to the WIPO IPC-Technology Concordance Table (see SCHMOCH 2008: 9-10 or section 5.3.4) 35 technological fields are distinguished.

3.3.2 Data collection

Other than the data presented in section 2, the data necessary for the social network analysis does not need to be collected from diverse sources, but is instead retrievable from *one* source (assuring a high degree of consistency and completeness): the EPO Worldwide Statistical Patent Database version October 2007 (PATSTAT VERSION 10/ 2007), published on a regular basis by the Patent Statistic Task Force. Since nanotechnology is not depicted in a specific IPC-class until 2006 and the introduction of the nano-specific tag "Y01N" into PATSTAT was still in progress by 2007 (see HULLMANN AND FRYČEK 2007: 11), patent applications are searched for following the search strategy developed by Fraunhofer, ISI (in: NOYONS ET AL. 2003: 100-101, see section 5.2.2).

In other words, in section 3 (this section), the definition of nanotechnology is determined by the search strategy of Fraunhofer, ISI (in: NOYONS ET AL. 2003: 100-101, see section 5.2.2). The search strategy comprises a list of search strings which contain a certain combination of keywords. One example for such a search string is

> "S (NANOPARTICL? OR NANO(W)PARTICL?) NOT (ABSORB? OR INK OR POLISH?)" (NOYONS ET AL. 2003: 100)

In the string (indicated by the letter "S"), the punctuation "?" stands for "unlimited truncation (0 or any number) and "W" stands for "directly adjacent terms in order specified" (see NOYONS ET AL. 2003: 101). Concretely, in the example, the data is searched for entries containing the word "nanoparticle" (or slight variations of it such as "nanoparticles", "nano particle" or "nano particles") but the word nanoparticle should not occur together with (variations of) the words "absorb", "ink" or "polish". The strings are transformed into SQL-code and run over the abstracts in the PATSTAT (VERSION 10/ 2007) database.

3.3.3 Data preparation: establishment of dataset 2

Having retrieved the patent applications, the data is merged into EXCEL-format. This dataset is further referred to as dataset 2. It is the data basis underlying this section. The subsequent paragraphs and descriptive statistics provide detailed information on dataset 2 to ease the interpretation of results. Having prepared the data in EXCEL spreadsheets, the data is transformed into a two-mode sociomatrix and then into an adjacency matrix using Java programming[89, 90]. The adjacency matrix is brought into a specific format so it can be processed by Pajek[91].

3.3.4 Data evaluation and descriptive statistics

(1) Application authority

The application authority is the patent office at which the patent is applied for. To secure data consistency, in the database underlying section 3 (database 2), only the *EPO and WIPO* are considered as application authorities. Patents applied for at these organizations are recognized as patents with a high technologic and economic impact (see FRIETSCH ET AL. 2008).

(2) Applicants sequence number/ Inventor sequence number

In PATSTAT (VERSION 10/ 2007), applicants are identified over the applicant sequence number: as patent applications can have more than one applicant, this number yields the position of an applicant on a patent application. Similarly, the inventor sequence number yields the position of the inventor. As in this book the focus is on companies which usually are applicants and not inventors, the data is restricted to applicants *(applt_seq_nr>0)*.

(3) Person country code

The person country code depicts the country an applicant or inventor is from. As in case of this book, domestic companies are of interest, the analysis is limited to Germany (entries with *person country code=DE*).

(4) Person name

In PATSTAT (version 10/ 2007), the person name is either the name of the applicant or of the inventor. As mentioned, in this book, the focus is on applicants. Applicants may be companies, but also, for example, persons. To distinguish companies from other types of applicants, person names need to be assigned to

---

[89] For a more detailed description of an adjacency matrix and a two-mode sociomatrix see CANTNER AND GRAF (2006: 466-467).

[90] See section 5.3.2 for the Java programs.

[91] Pajek: professional computer software for network analysis and visualization (see DE NOOY ET AL. 2005: XXIII).

categories. In a first step entries in the field “person name” are automatically assigned to one of the following groups: “company”, “academic R&D”, “non-university R&D”, “charity” or “person”. Table 22 (below) presents how the automatic assignment to the categories is achieved.

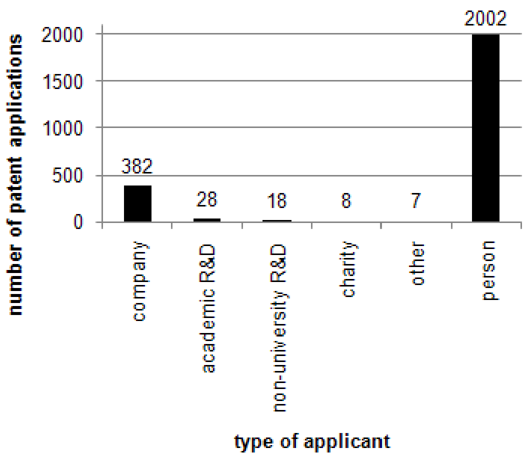
Table 22: Search strategy for the identification of applicant types.

Type of applicant	Search term
company	"GbR", "OHG", "kg", "gmbh", "MBH", "M.B.H", " ag", "aktiengesellschaft", "Genossenschaft", "LTD", "INC."
academic R&D	"HOCHSCHULE", "Universit"
non-university R&D	"FRAUNHOFER", "MAX-PLANCK", "Helmholtz", "LEIBNIZ", "MAX-DELBRUECK"
charity	“Stiftung”
person	if none of the above options are true
other	see remarks below

Source: Own determination.

The data is then *manually* reviewed to assure the correct assignment to the categories. The category “other” is chosen for entries which could not be assigned to one of the above groups. A manual unification of the company name follows as sometimes, the same company name occurs in different forms of spelling. Companies which are contained in the database with more than one legal form (e.g. AG and GmbH), are treated separately to avoid wrong assignments. Unfortunately, while possible for persons, for companies, the unification process cannot be achieved over the field “person_id” (in PATSTAT “person_id” is the key for the unambiguous identification of a person name) as companies usually have more than one person ID. Figure 16 (below) depicts the distribution of categories in the database.

Figure 16: Distribution of the types of applicants.



Source: Own *computations based on dataset 2*[92].

As the distribution visualizes, the share of *persons* amongst the applicants is highest. With 382 entries the second largest category is the category *company*. The large number of persons in the dataset occurs due to the fact that on patent applications, the inventors are often also listed as applicants. Additionally, it is likely that for example, on a patent application with six applicants there are five persons, but only one firm listed as an applicant. The reverse case of one person being listed on a patent application together with a larger number of companies does usually not occur (in the database, in altogether 18 of 1284 (~1.402%) patent applications, cooperations between two or more companies occur, see section 5.3.3). This commonly found ratio between persons and companies yields another explanation for the distribution depicted in Figure 16. The descriptive statistics presented in the following focus on those patent applications deriving from the *382 companies of database 2*.

(5) Priority date

Patents can be applied for at several patent offices. If so, the application date of the earliest application of a protective right can be used for an application at another patent office. In this case, the application date of the earliest application is considered the *priority date*.[93] Timely seen, the priority date is closest to the actual invention. Furthermore, compared to other dates possibly included on a patent application, for example the publication date, the priority date is clearly and unambiguously defined.

[92] See section 3.3.3.
[93] See Deutsches Patent- und Markenamt: Glossar, http://www.deutsches-patentamt.de/service/glossar/n_r/index.html#a1, 10 December 2010.

In this work, since PATSTAT (VERSION 10/ 2007) is used – implying incomplete data for the years 2006 and 2007 – only nanotechnology patent applications with priority years between [1978, 2005] are considered. The lower limit, 1978, depicts the year where the first nanotechnology patent application is filed in dataset 2. The upper limit of 2005 corresponds to the maximum complete year in the sample. Accordingly, the period of analysis comprises 28 years. In preparation of the social network analysis, cohorts are introduced (see Table 23 below).

Table 23: Distribution of companies over the cohorts.

Cohort	Priority date	Companies
1	[1978, 1985]	18
2	[1986, 1990]	41
3	[1991, 1995]	72
4	[1996, 2000]	125
5	[2001, 2005]	231

Source: Own computations based on dataset 2.

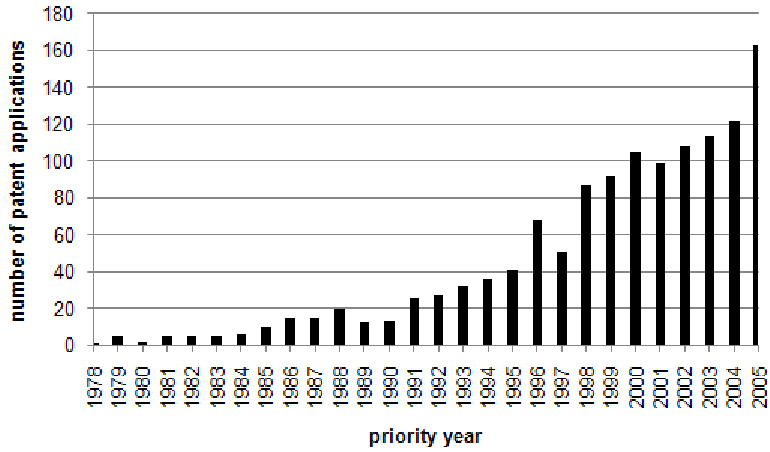
With the exception of the first cohort which comprises eight instead of five years, the cohorts are of equal length.[94] Since a company may be active in diverse cohorts, in sum, the number of companies exceeds the actual number of different companies (382 companies) in dataset 2. In general, the number of companies within the cohorts grows steadily, which confirms the picture of increased foundries drawn in section 2 (see p. 26, Figure 5). This might be regarded as a sign of nanotechnology becoming increasingly important over time. As the subsequent paragraph (6) explains though, this development should be seen in relation to the overall patenting behavior at the EPO.

(6) Application ID

The application ID allows for the unique identification of a patent application. Figure 17 (below) depicts the development of patent applications of domestic companies per priority year.

[94] The complete frame of analysis is not cut to 25 years or less to secure having a maximum interval to analyze. Furthermore, the first period is chosen to be longer as the number of firms in this time span is comparably small.

Figure 17: Timely development of patent applications.



Source: *Own computations based on dataset 2.*

Figure 17 reveals that the number of patent applications increases rapidly between 1978 and 2005. As in case of the previously depicted growing number of companies over time, this could be captured as a sign of increased popularity of research in the area of nanotechnology. This picture is valid, but needs to be slightly relativized because in 1978, the first applicants filed for a European patent protection at all.[95] The first couple of years following the foundry could therefore be regarded as "years of approval". Independent of nanotechnology, afterwards, a *general* notion of increased patenting activities is perceivable. To minimize the risk of observing a *general* tendency of growing patent applications instead of observing a nanotechnology-specific trend, the growth rates of *nanotechnology* patent applications (deriving from domestic nanotechnology companies) are held against the growth rates of *all* domestic patent applications at the EPO. As can be computed from the sample data, the average growth rates of *nanotechnology* patent applications is 1.167 in the relevant time span between 1991 and 2005. This growth rate merely reflects the growth rate of nanotechnology patent applications deriving from *domestic companies*; other applicant types are not considered. In the same period of time, by contrast, the growth rate of *all* domestic patent applications is comparably smaller (1.057)[96]. It therefore seems justified to assume that the observed tendency of growing patent applications in the area of nanotechnology is not merely the result of growing patenting activities in general. In total, *1284 nano patent applications* are captured in the period of analysis.

[95] See EPO: History, http://www.epo.org/about-us/office/history.html, 21 December 2010.
[96] To compute the average growth rate of *all* domestic patent applications at the EPO, data from NEUHÄUSLER (2008: 21) is used. In his work, NEUHÄUSLER (2008: 21) presents a table in which he depicts the absolute number of patent applications at the EPO of several countries (amongst other Germany) in the time span between 1991 and 2005. He states to have derived these numbers from EPPATENT and WOPATENT.

(7) Technological fields

Each of the analyzed patent applications contains information on the IPC-class(es) it touches upon, e.g. F41H 9/06. The first letter (“F” in the example) refers to the section. Table 24 (below) depicts the eight IPC-sections.

Table 24: IPC-sections.

Section Symbol	Section Title
A	HUMAN NECESSITIES
B	PERFORMING OPERATIONS; TRANSPORTING
C	CHEMISTRY; METALLURGY
D	TEXTILES; PAPER
E	FIXED CONSTRUCTIONS
F	MECHANICAL ENGINEERING; LIGHTING; HEATING; WEAPONS; BLASTING
G	PHYSICS
H	ELECTRICITY

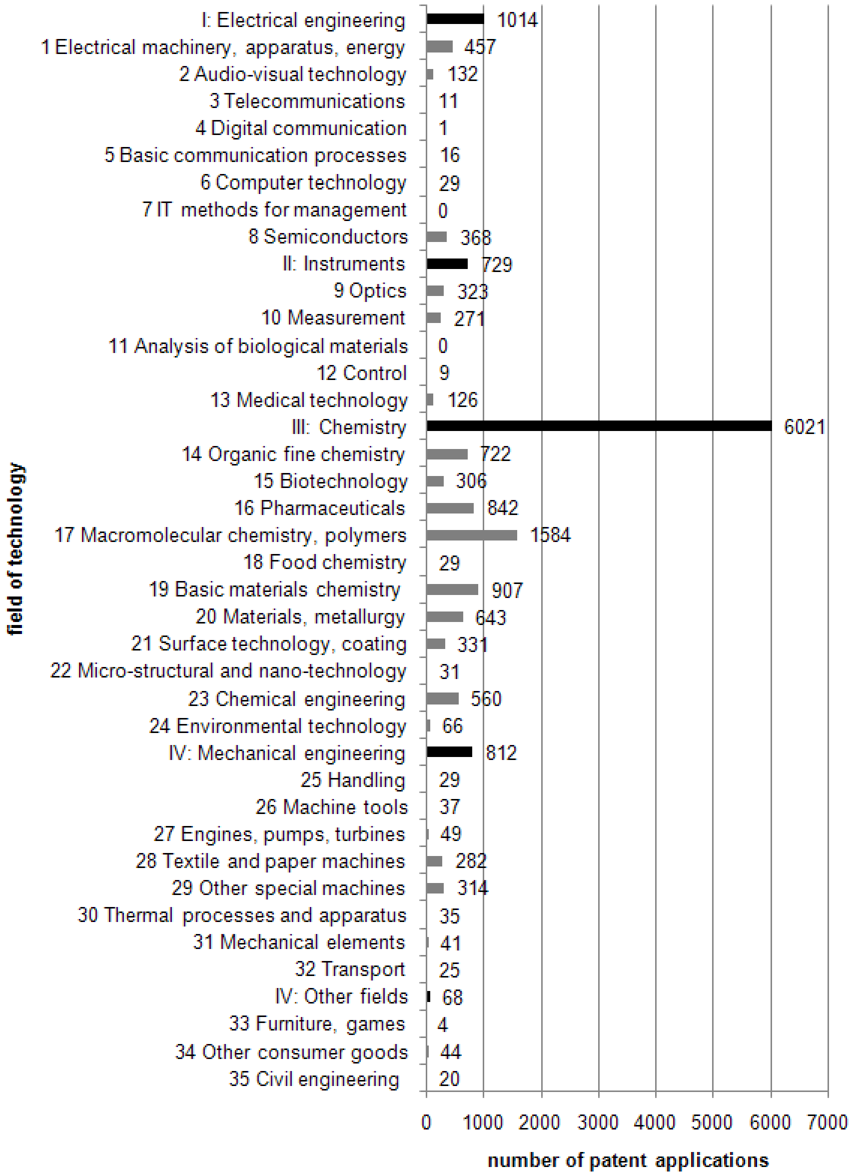
Source: WORLD INTELLECTUAL PROPERTY ORGANIZATION (2009: 3).

The two-digit number following the first letter (“41” in the example) denotes the class, the second letter (“H” in the example) demarks the subclass etc.[97] Since for the purpose of this book an analysis on IPC subclass level (e.g. F41H) or even finer is too subtle – resulting into a too large number of unconnected nodes in the network – and furthermore, because technological areas are difficult to perceive, an analysis on this level is not aimed at (depending on the evolutionary path of nanotechnologies' technological fields, the finer distinction might be attractive to use in future works though). Instead, in this book, the classification scheme as recommended by SCHMOCH 2008 is employed: In order to ease country comparisons, the author suggests a technology classification based on IPC codes. The classification distinguishes amongst 35 fields of technology (see SCHMOCH 2008: 9-10 or section 5.3.4). Figure 18 (below) depicts the distribution of the retrieved nanotechnology patent applications to the five i.e. 35 technological fields in the database.[98] In the figure, the black bars depict the main fields and the grey bars below depict the fields of technology assigned to the respective main field.

[97] See WORLD INTELLECTUAL PROPERTY ORGANIZATION (2009: 3-5).

[98] For three subsections (A61P, B01D or H04N) a corresponding field of technology could not be found. Companies which applied for patents in these fields are not depicted in the statistics presented before and are not included in the social network analysis.

Figure 18: Distribution of patent applications to the 35 fields of technology on the basis of the WIPO IPC-Technology Concordance Table.



Source: Own computations based on dataset 2 and on WIPO IPC-TECHNOLOGY CONCORDANCE TABLE.

With the exception of two fields (7 IT methods for management and 11 Analysis of biological materials) all other fields of technology are covered – each to a different extent. Obviously, the technological area "Chemistry" is represented strongest in the database indicating a strong relation between nanotechnology and chemistry. This seems to confirm the impression of section 2 where the chemical industry also is amongst the top ten areas mostly referred to. While on the one hand the close relation to chemistry could be a mere consequence of the general industrial orientation of Germany, on the other hand it might be regarded as an indice of the general orientation of *nanotechnology* in Germany: Apparently, in Germany nanotechnology seems to be rather bottom up than top down-driven (also see p. 4, Figure 1). Within the area “Chemistry”, field 17 (Macromolecular chemistry, polymers) stands out. The technological field “22 Micro-structural and nano-technology” in which nanotechnology is explicitly mentioned is comparably small which could either suggest an imprecisely transformed searching procedure, an imprecise formulation of technological fields or again demonstrate the difficulty of restricting nanotechnology to a specific field of technology.

3.4 Methodological approach

In this section, the methodological approach applied in connection with the social network analysis is briefly sketched. The corresponding results are presented in section 3.5.

According to DE NOOY ET AL. (2005: 5) “[t]he *main goal* of social network analysis is detecting and interpreting patterns of social ties among actors.” As pointed out by WASSERMANN AND FAUST (1994: 4) these ties are assigned an essential role:

> “[...] [S]ocial network analysis is based on an assumption of the importance of relationships among interacting units. [...] relations defined by linkages among units are a fundamental component of network theories.”

Optically, a network consists of vertices and lines depicting actors and the connections between the actors[99]. As mentioned in section 3.2, in this book, based on nanotechnology patent applications, networks of technological overlap are constructed for five subsequent cohorts (altogether comprising the time between 1978 and 2005) following CANTNER AND GRAF (2006). In the networks of technological overlap, the actors are domestic companies applying for one or more nanotechnology patent(s). Connections or ties between companies demark a present technological overlap between them. Following CANTNER AND GRAF (2006: 466), they emerge whenever two companies apply for a patent in the same technological class. This implies that in each network of technological overlap, all companies applying for a patent in the same technological field are connected to each other. Altogether, 35 technological fields are distinguished following the WIPO IPC-Technology Concordance Table (see SCHMOCH 2008: 9-10 or section 5.3.4). Since ties between companies result from a present technological overlap, the connections between companies are of *potential* and (at least mostly) not of *actual* nature (e.g. in the sense of a cooperation). At most, they might

[99] In the methodological part depicted in section 3.4, the general term “vertex” or its synonym “actor” is used when referring to points in the network. In the remainder of section 3, these terms stand for companies.

be regarded as an *indicator* for explaining possibly occurring cooperations. In this book, the networks of technological overlap are used to explain the technological survival of companies in the area of nanotechnology. In order to construct such networks, it is necessary to generate an adjacency matrix on the basis of a two-mode sociomatrix.[100]

3.4.1 General network structure

Descriptive statistics yield information on the general network structure: For instance, the *number of actors* gives some information on the network size. In preparation of the analysis, in this book, the number or set of actors (per cohort) is devided into *exiting* companies[101] and *permanent* companies[102]. Certainly, companies can also enter into cohorts (*entrants*[103]). However, since in this book, the *survival* patterns are in focus, entries are of secondary relevance and therefore, in the following, are only included for reasons of completion.

While the number of actors is useful for tracking the development of the network size, it does not yield any information on *how* the actors are organized. In an extreme case the actors might not be connected to each other at all; in the other extreme case they might all be connected to each other (for instance over technological fields).

To find out about the organization of the actors, it is useful to obtain knowledge about the components within a network. According to DE NOOY ET AL. (2005: 318), "[a] […] component is a maximal […] connected subnetwork." Simply put, a component is a "separate" subnetwork within a network. In terms of learning about the organization of actors, in particular, the *number of components* (stating how many separate entities or disconnected parts there are in the network) is helpful to know. If there are equally as many actors in the network as there are components, the network only consists of isolates (the smallest possible component size). If there are fewer components than actors, obviously (at least some of) the actors are connected to each other. In this regard, the *number of isolates* yields the number of isolated actors in the network and, respectively, the *size of the largest component* yields the component including the most actors.

However, though both, the number and size of components, are useful for obtaining information on the organization of the network, they are not useful in terms of finding out about the *strength* of connection between the actors. For this kind of

[100] For a more detailed description of an adjacency matrix and a two-mode sociomatrix see CANTNER AND GRAF (2006: 466-467).

[101] A company obtains the status "exiting company (from cohort t)", if it holds a patent application in cohort t, but not in cohort (t+1).

[102] A company obtains the status "permanent actor" (in cohort t) if it applies for a patent in cohort t and in cohort (t+1).

[103] A company obtains the status "entrant (to cohort (t+1))", if it has not applied for a patent in t, but applies for a patent in (t+1).

evaluation the measures *density* and *degree* are useful. According to DE NOOY ET AL. (2005: 319) they are defined as follows:

> "Density is the number of lines in a simple network, expressed as a proportion of the maximum possible number of lines."

The *degree* on the other hand yields the number of lines incident with a vertex (see DE NOOY ET AL. 2005: 319). Consequently, the *average degree* gives information on the mean number of lines incident with a vertex of a network.

All previously described measures help gaining an insight into the network in general. However, as in case of this book, it might also be necessary to learn about the *roles* that certain actors play within the network. In social network analysis, one typical approach to do so is to locate *central* and *peripheral* actors. With regard to the social network analysis performed in section 3.5, the subsequent paragraph describes commonly used measurement concepts to identify central and peripheral actors within a network.

3.4.2 Identification and implication of (de-)centralized networks

According to JANSEN (2006: 127) *centrality* and *prestige* are two network analytical concepts asking for the importance, public visibility or "prominence" of actors. Concepts of *centralization* assume *undirected relations* between actors[104] while *prestige* assumes *directed relations*. Since in this work there are only undirected networks, the focus is on concepts of centralization.

DE NOOY ET AL. (2005: 125-134) suggest three basic measurement concepts for determining centralization: degree-, closeness- and betweenness-based measurement concepts. Within each measurement concept, two perspectives are distinguished: the actor-based view (the centrality of a singular actor within a network) and the network-based view (the centrality of an entire network). In the following, in line with DE NOOY ET AL. (2005: 123), the term "centrality" is used for measures referring to positions of individual vertices within a network and the term "centralization" is used for measures characterizing an entire network. All measurement concepts (degree-, closeness- and betweenness-based measurement concepts) have in common that – in order to obtain network centralization – information on actor centrality needs to be present.

In brief, prior to further methodological remarks presented in later sections, the degree-based measurement concept focuses on direct connections between actors. The closeness-based measurement concept also takes indirect connections between actors into account. According to DE NOOY ET AL. (2005: 131), degree and closeness centrality are based on the reachability of a person within a network: Both approaches assess how easily information may reach a person. Diffe-

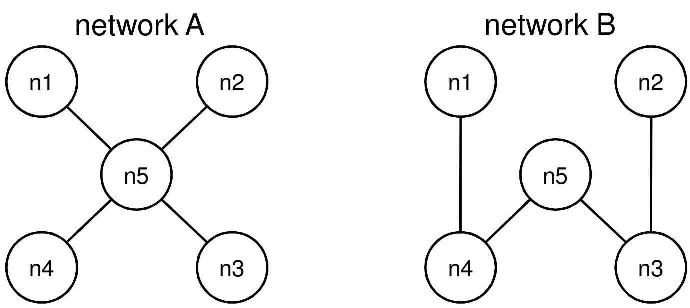
[104] Concepts of centrality were originally developed for undirected relations. The centrality of actors in the sense of their participation at activities is also computable for networks with directed asymmetric relations (see JANSEN 2006: 127).

rently put, JANSEN (2006: 135) says that degree and closeness centrality measure the independence of an actor from other actors as intermediaries: A central actor has many direct connections or indirect paths to all others and therefore is not or only seldom dependent on others. Finally, the betweenness-based measurement concept assesses in how far an actor has an intermediary position in the network. Following DE NOOY ET AL. (2005: 131), this approach rests on the idea of a person being more central if he or she is more important as an intermediary in the communication network.

As the three measurement concepts place their emphasis on different aspects, the question is which measurement concept should best be used in the context of the analysis aimed at in section 3. Since the focus of section 3 is to explain the technological survival of companies by their *direct* technological overlap (technological overlap thereby resulting from the choice of the companies' technological fields and therefore to be perceived as a measure of potential interaction), of all three measurement concepts, in this book, the focus is on the degree-based measurement concept. The closeness- and the betweenness-based measurement concept are regarded as less adequate for the detection of central and peripheral actors, as their explanatory power is more adequate for networks where the connections between actors are *actual*; in networks of technological overlap, they do not come along with a natural interpretation.

For reasons of completion, comparability and to provide an alternative view, in this book, *all* measurement concepts are explained. The degree-based measurement concept is presented in section 3.4.2.1. Due to their assigned subordinate role, the description of the closeness- and the betweenness-based measurement concepts is moved to the appendix (see section 5.3.5). To ease understanding, in addition to the methodological remarks on each measurement concept, exemplary calculations are included based on the networks shown in Figure 19 (below). The example is adopted from DE NOOY ET AL. (2005: 125).

Figure 19: Exemplary networks.



Source: DE NOOY ET AL. (2005: 125).

In brief, network A shows a star network[105], network B depicts a line network. In the following, after the methodological remarks and the example, a brief interpretation of the respective measurement concept including its advantages and disadvantages is given.

Section 3.4.2.2 then focuses on the implications that can be derived when ascertaining that a network is central or, respectively, decentral. Possible advantages and disadvantages of each are highlighted.

3.4.2.1 Degree-based measurement concept

The degree-based measurement concept comprises two actor-based measures (the degree centrality and the normalized degree centrality) and one network-based measure (degree centralization). These measures are presented consecutively.

The *degree centrality* of actor n_i ($C_D(n_i)$) captures the number of direct connections to other actors (see JANSEN 2006: 137). Mathematically, $C_D(n_i)$ is computed according to the formula (see JANSEN 2006: 137)

$$C_D(n_i) = \sum_j x_{ij} \text{ and } i \neq j$$

where x_{ij} demarks the presence of a direct connection between actor n_i and actor n_j. If a connection between actors n_i and n_j is present x_{ij}=1, if there is no connection between the actors, x_{ij}=0 (see FREEMAN 1979: 220).

In an undirected network without multiple lines or loops (which is the case in this book) and the network consisting of altogether n actors, at maximum, an actor may be connected to all (n-1) other actors in the network. At minimum, it may have no connections to other actors. In the example network A (see p. 81, Figure 19) n_5 is connected to all other actors. It therefore has a maximum degree centrality of $C_D(n_5)$=(5-1)=4. The remaining actors have a degree centrality of $C_D(n_1)=C_D(n_2)=C_D(n_3)=C_D(n_4)=1$. In network B, $C_D(n_1)=C_D(n_2)=1$ and $C_D(n_3)=C_D(n_4)=C_D(n_5)=2$. Altogether, $C_D(n_i)$ ranges in the interval [0, (n-1)]. As the upper boundary shows, the degree centrality is dependent on the network size. When networks of different sizes are present, this makes comparisons between actors of these different networks difficult.

To neutralize the effects of different network sizes, $C_D(n_i)$ is related to its maximal possible value of (n-1) (see JANSEN 2006: 133, 137). This measure is then referred to as the *normalized degree centrality* of actor n_i ($C_{nD}(n_i)$). Mathematically, this implies (see JANSEN 2006: 137)

[105] General definition: "A star-network is a network in which one vertex is connected to all other vertices but these vertices are not connected among themselves." (DE NOOY ET AL. 2005: 324).

$$C_{nD}(n_i) = \frac{C_D(n_i)}{n-1}$$

It ranges in the interval [0, 1]. In example network A (see p. 79, Figure 19) $C_{nD}(n_5)=1$. The remaining actors have a normalized degree centrality of $C_{nD}(n_1)=C_{nD}(n_2)=C_{nD}(n_3)=C_{nD}(n_4)=1/4$. In network B, $C_{nD}(n_1)=C_{nD}(n_2)=1/4$ and $C_{nD}(n_3)=C_{nD}(n_4)=C_{nD}(n_5)=1/2$.

Based on the *degree centrality* of singular vertices, it is possible to compute a measure to evaluate the degree centrality of an entire network, the so-called *degree centralization* (C_D). DE NOOY ET AL. (2005: 126) describe the degree centralization of a network as follows:

> "**Degree centralization** of a network is the variation[106] in the degrees of vertices divided by the maximum degree variation which is possible in a network of the same size."

Mathematically, C_D can be defined as (see JANSEN 2006: 139)

$$C_D = \frac{\sum_{i=1}^{n}[C_D(n^*) - C_D(n_i)]}{\max\sum_{i=1}^{n}[C_D(n^*) - C_D(n_i)]} = \frac{\sum_{i=1}^{n}[C_D(n^*) - C_D(n_i)]}{n^2 - 3n + 2}$$

where $C_D(n^*)$ is the degree centrality of the most central actor in the network and $C_D(n_i)$ is the degree centrality of actor n_i. The numerator is the variation in the degree of the actors. The expression in the denominator stands for the maximum possible sum of differences between the most central actor and all others in a network of size n which according to JANSEN (2006: 139) is achieved by the star network. In a star network, $C_D(n_i)=1$ pertains to all actors except for the central actor, who has a degree centrality of $C_D(n^*)=n-1$. Therefore:

$$\max\sum_{i=1}^{n}[C_D(n^*) - C_D(n_i)]$$

$$= (n-1) * [(n-1) - 1] + 1 * [(n-1) - (n-1)] = n^2 - 3n + 2$$

Since the observed network is always related to the maximum possible network centralization, C_D ranges in the interval [0, 1].[107] A low value implies that – according to the degree-based measure – a decentral network is present and a

[106] "Variation is the summed (absolute) differences between the centrality scores of the vertices and the maximum centrality score among them. [...]" (DE NOOY ET AL. 2005: 126).
[107] If the analyzed network contains multiple lines or loops, DE NOOY ET AL. (2005: 126) portend that the degree of a vertex is not equal to the number of its neighbours which may lead to a degree centralization of greater than one. In this case they advice not using this measure. This however, is not relevant in terms of the application presented in this book.

high value implies that a rather centralized network is present. In example network A (see p. 81, Figure 19) C_D=1, in network B, C_D=1/6.

As the measure “degree centrality” is based on the number of lines incident with a vertex, one essential advantage is that the measure can easily be calculated for every actor in the network – no matter whether it is isolated from the others or is part of a component. Since the degree centrality is always computable for each actor of a network, also the normalized degree centrality and the degree centralization are always computable. This eases comparison between actors and networks: Actors with a higher number of incident lines may be classified e.g. as more active in communication or as more technologically diversified than those with fewer incident lines.

On the other hand, the fact that only direct neighbors of a vertex are taken into account may be unsuitable when the focus of the analysis is another than in this book, for example when assessing in how far an actor is independent from other network members or when being primarily concerned with the ability of an actor to control communication within a network. If these aspects are central to the analysis, closeness- and betweenness-based measurement concepts[108] should be considered.

Having presented the degree-based measurement concept for determining centrality and centralization in the networks of technological overlap, the subsequent paragraph is concerned with the implications that can be derived when finding out that a network is central or decentral.

3.4.2.2 Implications of central and decentral networks

In many cases, social network analysis refers to the analysis of networks depicting “actual” connections, for example networks depicting cooperations or networks depicting personal connections within corporate divisions. However, networks need not to be organized intentionally: The networks of technological overlap for example (which are analyzed in this book) depict overlaps in the choice of technological fields of companies; they merely depict the “potential” of cooperation (see CANTNER AND GRAF 2006: 466). Depending on the kind of network, implications of central and decentral networks need to be adapted. Aiming at presenting implications of central and decentral networks, this section is structured as follows: in a first subsection, implications of central networks are presented for networks depicting actual connections and for the networks of technological overlap. In a second subsection, implications of decentral networks are shown for both kinds of networks.

[108] See section 5.3.5.

(1) Implications of central networks

In an extreme case, in central networks there is one central actor having connections to all other network members while all other network members do not have a connection to each other. This description pertains to the previously described star network (see section 3.4.2).

In terms of “actual networks”, according to JANSEN (2006: 127, 138) network centralization is a measure for a groups’ capacity of solving problems. Following the author, the speed and efficiency of processing tasks, the satisfaction of group members and their perception of leadership as well as the groups’ ability to organize and to deal with conflicts are related to the tendency of an actor to be outstandingly central. Concretely, the assumption is that prominent actors have access to network resources, the ability to control and have access to information. The reverse side of the medal is that networks in which one actor controls the information flow to such an extent are rather error-prone: being the only one with access to all resources the central actor could (intentionally or unintentionally) withhold or manipulate information. Since none of the other network members are able to directly communicate with each other they cannot verify the obtained information.

Another characteristic of central networks is that – with the absence of the central actor – the entire network collapses in the sense that information cannot be transferred at all. While in this respect a central network is very fragile, on the other hand it allows for an easy integration of additional actors. This might be of advantage in specific networks, for instance in computer networks where further computers can simply be connected to the hub[109] permitting for a high flexibility.

In the networks of technological overlap, “central actors" have a high technological overlap to other companies. They have access to many technologies which could protect them from failing in two respects: Even in case of the unsuccessful pursuit of one technology there are always alternatives to follow up to. Furthermore, new options for inventions could arise due to the fact that they are able to discover interfaces. Additionally, the technological scope and inventiveness of the central actor may provide an ideal seedbed for cooperations: Assuming that each company’s technological focus is known to others, it is possible for the central actor to search for companies complementing its technological scope to for example further expand its position and to set standards. This could lead to a monopoly position of the central actor.

Contrary to actual networks, in the networks of technological overlap, the absence of the central actor does not imply the interruption of information flow: Since no actual information flow takes place, the information flow cannot be disturbed. The absence of the central actor would merely imply that the remaining

[109] See NETZWERKTOTAL DAS NETZWERK PORTAL: http://www.netzwerktotal.de/netzwerktopologie.htm, 19 January 2010.

actors would work in their own technological fields with no other company following up to the same technology.

(2) Implications of decentral networks

In decentral networks (in an extreme case) all actors are connected to each other. As this is the case, in "actual" networks, a monopoly position of a singular actor is not given. All actors have access to all resources and information can be spread by all actors in the network. This leads to a high data redundancy which allows each actor to verify the obtained information. Even if an actor breaks away, the remaining network stays intact; information may still flow. Therefore, a decentral network is very stable.

On the other hand, according to SCHEIDEGGER (2008: 504) (who analyzes the impact of structural holes on the success in career), group members with intensive connections amongst each other tend to homogeneity in views, opinions and behavior.[110] New information can hardly be created from the same pool of information.

Aside from the creation of knowledge, the high data redundancy might be rather time consuming as in networks, such as personal networks, each person listens to the same information many times. A coordinating instance is simply missing. Also, while central networks allow for the easy integration of additional actors, decentral networks are not equipped with this kind of flexibility. Connections between all actors need to be established first which – depending on the kind of network – can be difficult to achieve.

For the networks of technological overlap, in the first place, decentrality in an extreme case implies that all actors follow up to the same technologies or at least have a greater technological overlap in their pursued technological fields. This implies high competition, but also a stable technological ground: if one company fails, there are other companies following up to the technological fields chosen by the company which failed.

(3) Summary on implications of central and decentral networks

In summary, a general recommendation which kind of network – a central or a decentral network – is best cannot be given. Which organizational form is the most suitable one depends on the underlying data and, most importantly, on the perspective and general intention: From the perspective of an internal corporate

[110] According to SCHEIDEGGER (2008: 505), by contrast, a structural hole implies that contacts to various clusters are present and actors are integrated in various information flows. This enlarges the pool of alternative views and behavior. Information advantages result from the possibility of creating rich knowledge from selection and synthesis. A timely advantage because of an early perception of new information occurs. Due to recommendations of actors of other groups visibility is increased and the actor is at the right time at the right place. Both aspects are important for his successful career.

network (a network within a company) with the aim of quickly distributing information to all employees, a central network might be the best organizational form because it prevents actors to repeatedly listen to the same information (which costs time). From the perspective of an external corporate network (a network between companies) with the aim of distributing information to all companies without assigning one actor a central role, a decentral network may be most suitable for it secures competition. In the networks of technological overlap, a central network could imply that one actor increasingly gains power and influence. Decentral networks could imply high competition and a stable technological ground.

3.4.3 Discovering survival patterns

For the networks with a detectable center and periphery, it is possible to figure out whether the *position* of companies in the network (core or periphery) exerts an influence on the *status* or (in other words) *technological survival* (permanent/ exiting) of companies in the network. Concretely, the combined view of the *position* and *status* of companies in the networks can be viewed as the merging of distinct findings of section 3.4.1 and section 3.4.2. An analysis of this kind requires the thorough determination of core and peripheral actors in the networks in terms of their survival patterns.

In a first step, this can be achieved *graphically* by marking permanent and exiting actors in the networks of technological overlap and by then deducing statements concerning their survival. The graphical approach is useful for gaining a first impression of the survival patterns.

Since the graphical approach only allows for a subjective assessment, in a second step, an *analytical* approach can be added which allows distinguishing between core and peripheral actors in a more systematic manner. In this respect, the normalized degree centrality can serve as a classification criterion between core and peripheral actors. A sensitivity analysis may show in how far the choice of the "boundary normalized degree" centrality, i.e. the determination of core and peripheral companies, influences the survival patterns. Additionally, similar to the work of CANTNER AND GRAF (2006: 472), the connections within the group of permanent actors and within the group of exiting companies can be analyzed to give a possible explanation for the survival of firms.

In a supplementary investigation, in a third step, the technological focus areas of core and peripheral companies can be contrasted against each other to derive additional insights into possible reasons of survival.

3.5 In-depth analysis: network of technological overlap

In this section, five networks of technological overlap are constructed. All networks are submitted to several examinations to accept or reject the three hypotheses presented in the introductory paragraph of section 3. Finally, section 3.5 finishes with a summary and conclusion of the results.

3.5.1 Structural evolution

Hypothesis 4 concerns the structural evolution of the network. The assumption is that the network of technological overlap changes strongly over the diverse cohorts in terms of the actors which are part of the network and in terms of the ties between these actors. In order to accept or reject this hypothesis it is necessary to apply objective measures which help to assess the extent to which a structural change takes place. The measures presented in section 3.4.1 serve to give an insight on the structural evolution of the networks. Table 25 (below) summarizes the results for the five cohorts.

Table 25: Descriptive statistics of the networks of technological overlap[111].

	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5
# actors	18	41	72	125	231
# exits	*10*	*26*	*54*	*82*	
# permanent	*8*	*15*	*18*	*43*	
# entrants		33	57	107	188
# components[112]	6	2	1	2	1
# isolates	3	0	0	0	0
SLC[113]	11	39	72	123	231
average degree	3.111	11.073	23.944	41.968	64.840
density[114]	0.183	0.277	0.337	0.338	0.282

Source: Own computations based on dataset 2.

As Table 25 shows, while – with the exception of the first cohort – the cohorts are of equal length, the size of the network grows strongly in the observed time span between 1978 and 2005. Compared to the network of the first cohort, the network depicted in the fifth cohort contains around 12.833 times more *actors* (=companies). In general, the actors in each cohort can be distinguished into *exiting* and *permanent actors.*[115] For example, of the 18 actors in cohort 1, 10 experience a technological exit after cohort 1 and 8 are also actors in cohort 2. Of course, companies can also *enter* into cohorts. For example, in cohort 2, additional to the 8 actors from cohort 1, 33 new actors enter into the second cohort,

[111] Conform with section 2, the decimals in this section are rounded to three decimal places. For computations, the exact values are used.
[112] Minimum component size is one vertex.
[113] Size of largest component.
[114] No loops allowed.
[115] The exception is cohort 5 for which – since there is no cohort following cohort 5 – permanent and exiting actors cannot be determined.

resulting into the 41 actors in cohort 2. Of these, 26 exit after cohort 2 while 15 remain actors in the subsequent cohort 3 and so on.

It is striking that despite the continuous growth of the networks, the number of *component*s decreases from six components (amongst these three *isolate*d companies) in the first cohort to two or even one component(s) in the remaining cohorts (no isolates). As a consequence of the *decrease* in components and the *increase* of the number of actors, the size of the respective *largest component* grows strongly. In direct comparison to the first cohort, the largest component in the fifth cohort is exactly 21 times larger.

Likewise, the *average degree* of the vertices experiences a strong growth: While – on average – a company is connected to 3.111 companies in the first cohort, in the second cohort it is already connected to 11.073 and in the fifth cohort even to 64.840 companies. On first sight, the networks therefore seem to become increasingly connected. However, it needs to be kept in mind that in the first cohort, at maximum, a company may be connected to 17 other companies while in the fifth cohort, at maximum, a company may be connected to 230 other companies. Therefore, on average, the number of connections changes roughly from 3.111 connections/17 possible connections (=density of 0.183) in the first cohort to 64.840 connections/230 possible connections (=density of 0.282) in the fifths cohort. In general, the *density* of the network shows a slightly varying course: From the first to the fourth cohort, the density of the network increases indicating an increasing connection between the companies. The density then decreases between the fourth and fifth cohort. The observation of companies becoming increasingly connected to each other is therefore valid but a little relativized.

Figure 20 to Figure 24 (below) show the networks of technological overlap for the five cohorts. The network visualization is achieved with the program Pajek. Similar to CANTNER AND GRAF (2006: 467-468) in the networks, the *vertices* symbolize actors (=companies) and the *size of the vertices* depicts the number of nanotechnology patent applications of a company[116] in the respective cohort. A *line* depicts a technological overlap between two companies (it is important to note that companies may follow technologies within the area of nanotechnology which no one else works in; these are then not visible in the form of lines). The *line width* reflects the number of overlapping technologies. Further following CANTNER AND GRAF (2006: 467-468), to optimize *visualization*, only edges between companies are considered depicting a technological overlap in at least two technologies.

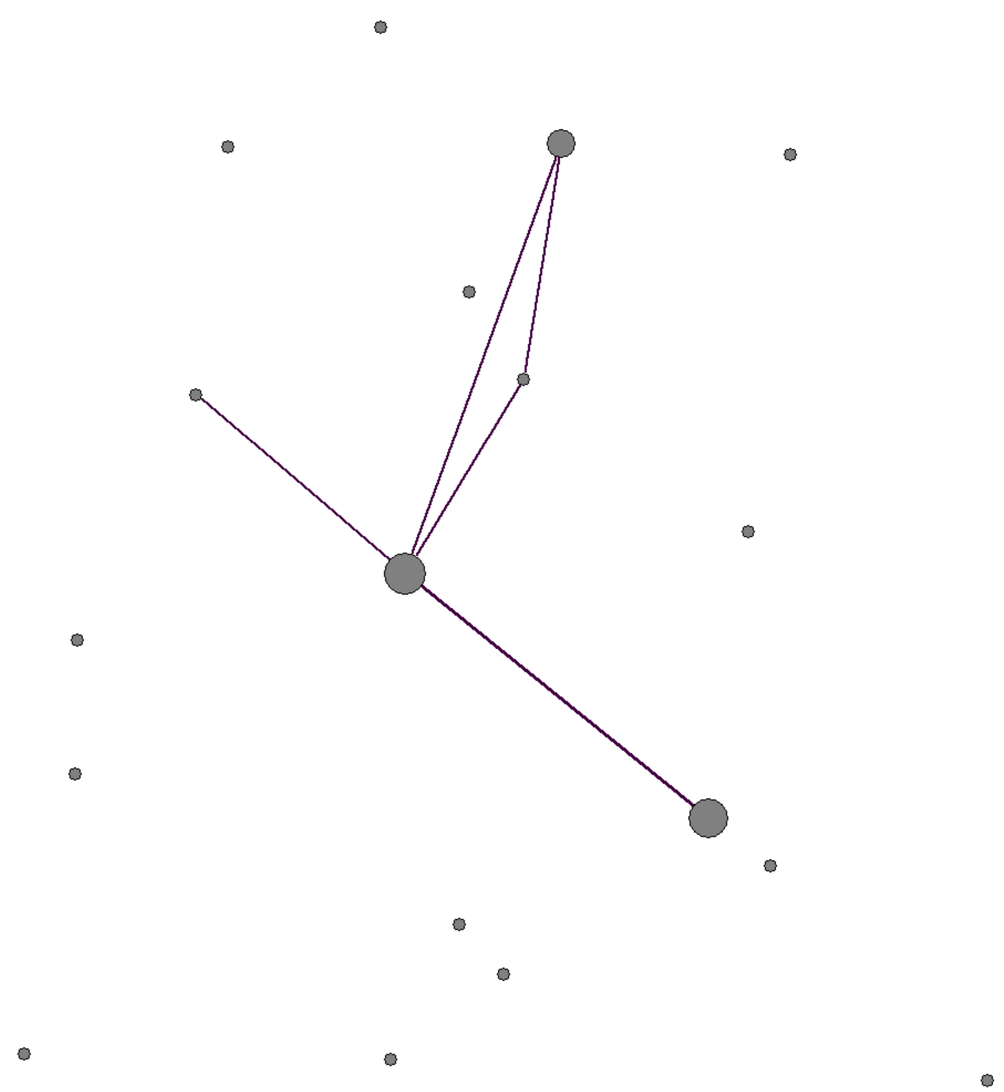
The figures are obtained using Kamada-Kawai and Fruchterman Reingold. Both are so-called "energy commands" responsible for the positioning of vertices in the networks. Kamada-Kawai produces regularly spaced results, especially for connected and not too large networks (see DE NOOY ET AL. 2005: 17). Fruchterman

[116] Note that the number of patent applications does not necessarily reflect the size of the company. It merely reflects its involvement into nanotechnology.

Reingold on the other hand separates unconnected parts of the network from another (see DE NOOY ET AL. 2005: 17).

Due to the large number of actors in cohort 4 and 5 the networks are very dense. Nevertheless, the figures are included to portray the general development of the networks.

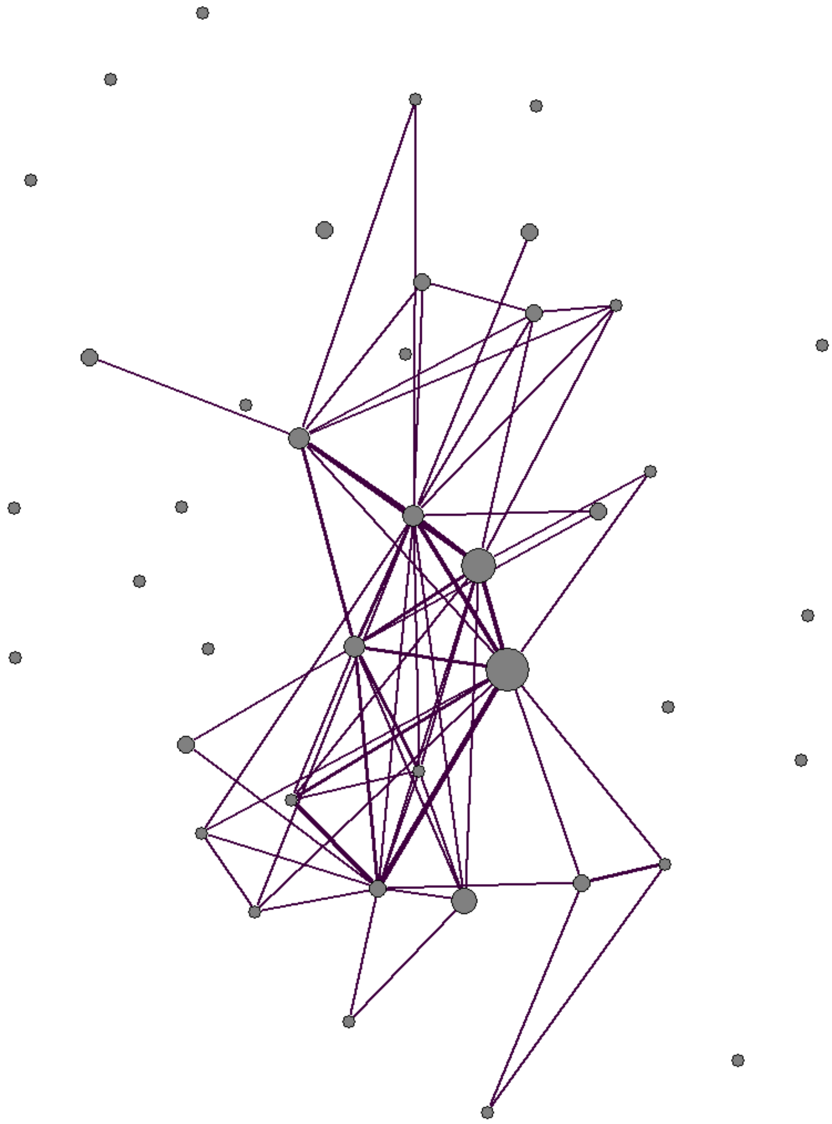
Figure 20: Network of technological overlap, cohort 1.



●	company
–	technological overlap

Source: *Own computations based on dataset 2.*

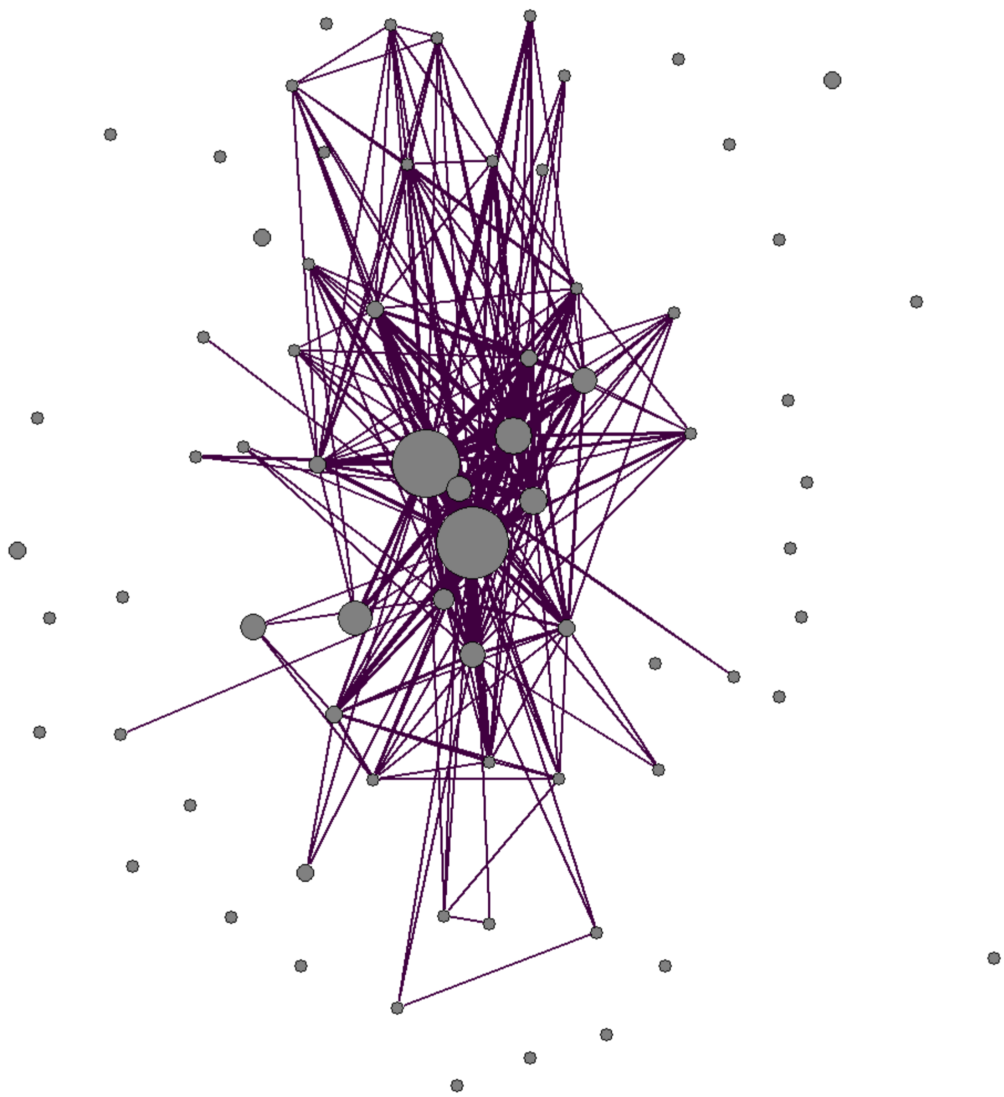
Figure 21: Network of technological overlap, cohort 2.



●	company
–	technological overlap

Source: Own computations based on dataset 2.

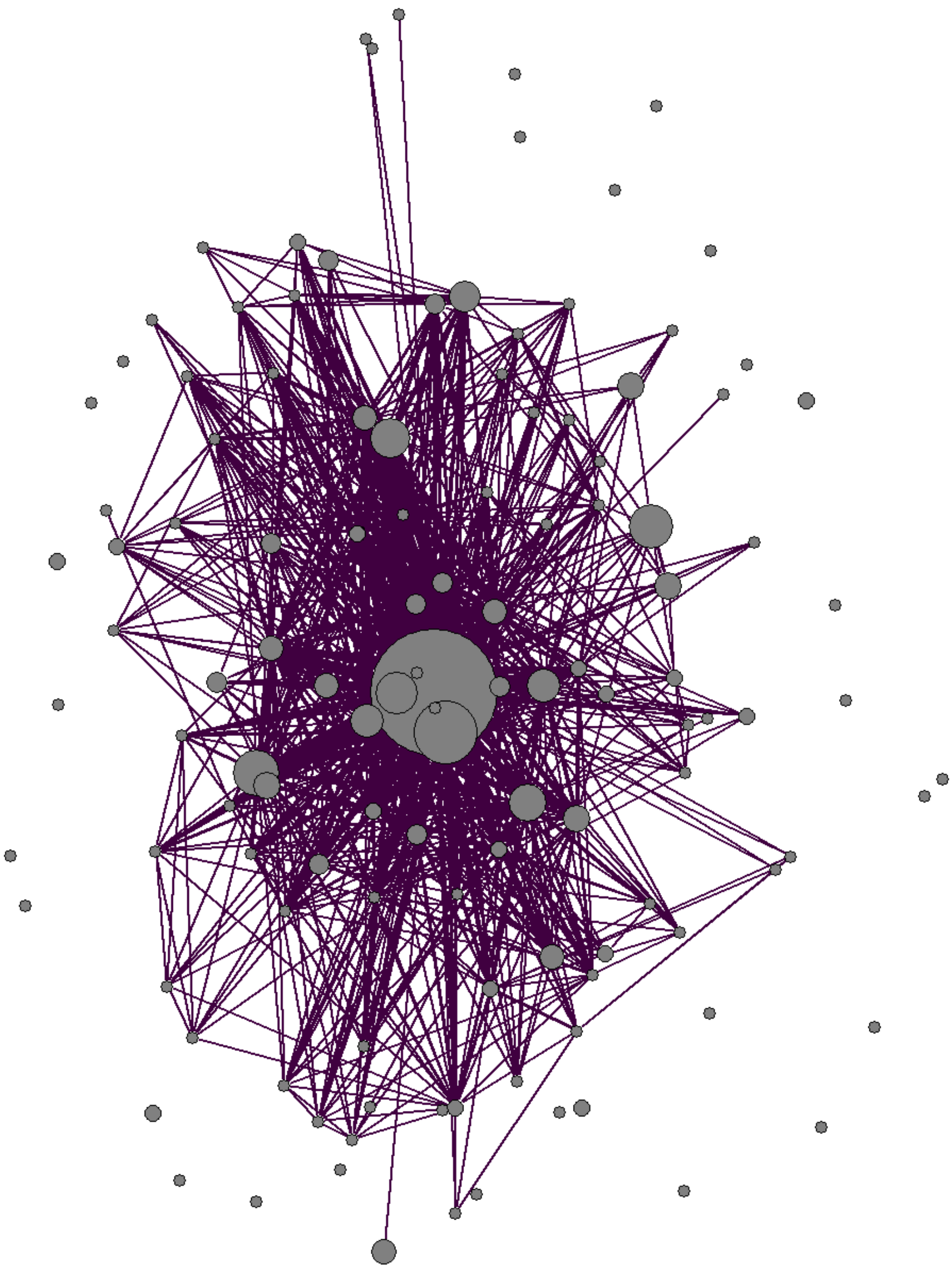
Figure 22: Network of technological overlap, cohort 3.



●	company
–	technological overlap

Source: Own computations based on dataset 2.

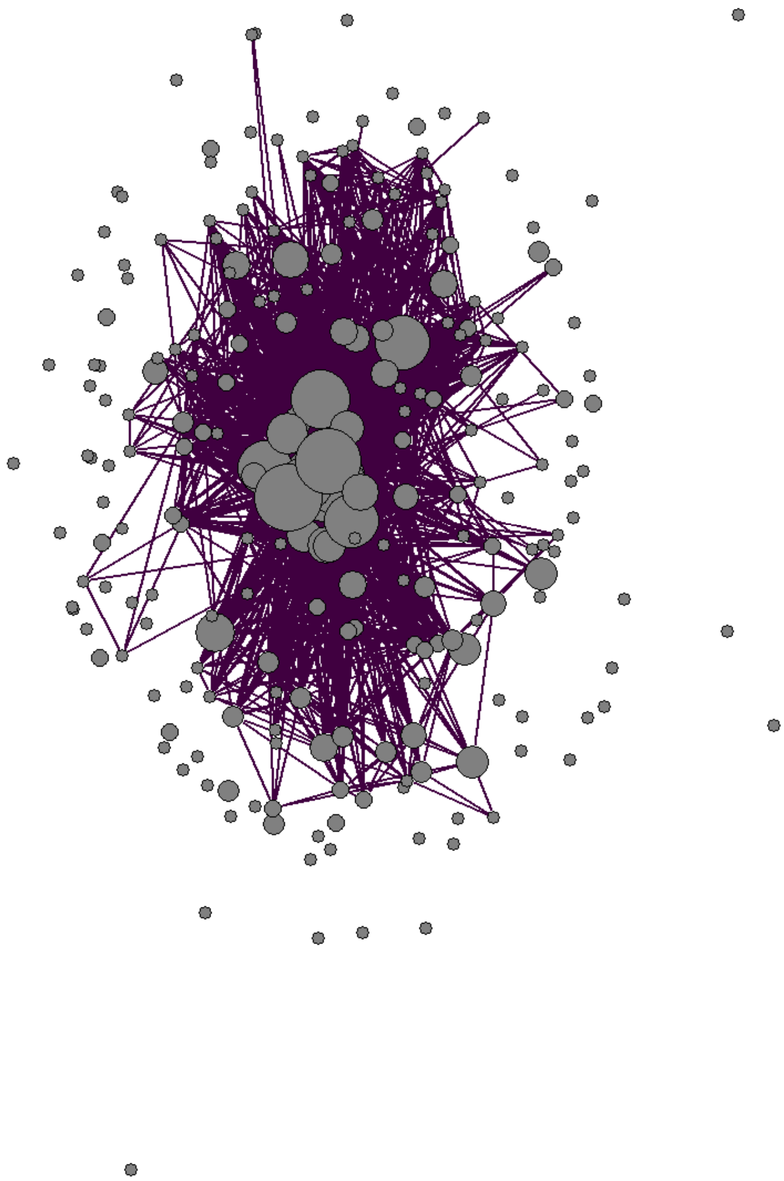
Figure 23: Network of technological overlap, cohort 4.



●	company
–	technological overlap

Source: Own *computations based on dataset 2.*

Figure 24: Network of technological overlap, cohort 5.



●	company
–	technological overlap

Source: Own *computations based on dataset 2.*

In general, the growing size of the networks of technological overlap (constructed by the help of nanotechnology patent applications) is very well depicted by the five figures. This concerns the number of actors as well as the ties between them. When looking at the underlying data material (not included in this work in this detail) it is striking that large chemical companies are always in the center of the networks. They always yield the majority of nanotechnology patent applications. This may partially explain the depicted pattern in Figure 18 (see p. 77).

With the gathered information on structural aspects available, hypothesis 4 can now be assessed:

> ***HYPOTHESIS 4:*** The structure of the network of technological overlap changes strongly over the observed cohorts in terms of the actors which are part of the network and in terms of the intensity of ties between the actors.

Regarding hypothesis 4, it can be depicted that the computed descriptive statistics as well as the visual impression indicate an enormous structural change of the network. While the companies with the majority of patent applications basically remain the same over all cohorts, in general, the strongly growing numbers of actors and the constantly high entry and exit rates (see p. 88, Table 25) indicate that in most other respects, the networks underlie high dynamics. For example, the network of cohort 1 no longer exists in its original form in the following cohorts. Merely some actors remain active in the following cohorts. The others exit from the network of technological overlap. Altogether, this seems to corroborate the enormous change and evolution that nanotechnology experiences. Concerning the ties between the actors, the average degree and density show that the companies are increasingly connected to each other. Hypothesis 4 may therefore be confirmed.

3.5.2 Technological centralization and decentralization

Several findings including network visualization point to companies playing different "roles" within the networks: Obviously, some companies seem to be more "central" to the networks than others. So far, this is rather an *impression* though; such "roles" are neither systematically defined nor identified yet. In preparation of hypothesis 5, section 3.5.2.1 is concerned with the definition and identification of such core and peripheral actors. Section 3.5.2.2 follows with conclusions.

3.5.2.1 Centrality in the networks of technological overlap

As outlined in section 3.4.2, due to the focus on the direct technological overlap of companies, the degree-based measurement concept is applied to assess centrality in the networks of technological overlap. Table 26 (below) summarizes the results of the normalized degree centrality and degree centralization. The normalized degree centrality is used instead of the degree centrality to ease comparison between actors of differently sized networks.

Table 26: Normalized degree centrality and degree centralization of the networks of technological overlap[117].

		Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5
Distribution of the normalized degree centralities	**(0.000, 0.200]**	10	14	17	30	84
	(0.200, 0.400]	7	22	33	55	101
	(0.400, 0.600]	1	1	12	26	30
	(0.600, 0.800]	0	4	7	10	12
	(0.800, 1.000]	0	0	3	4	4
Sum of actors		18	41	72	125	231
AM[118]		0.183	0.277	0.337	0.338	0.282
SD[119]		0.154	0.173	0.198	0.197	0.181
Interval […]		0.000, 0.529	0.025, 0.750	0.014, 0.944	0.008, 0.895	0.004, 0.878
C_D[120]		0.390	0.497	0.624	0.566	0.602

Source: Own computations based on dataset 2.

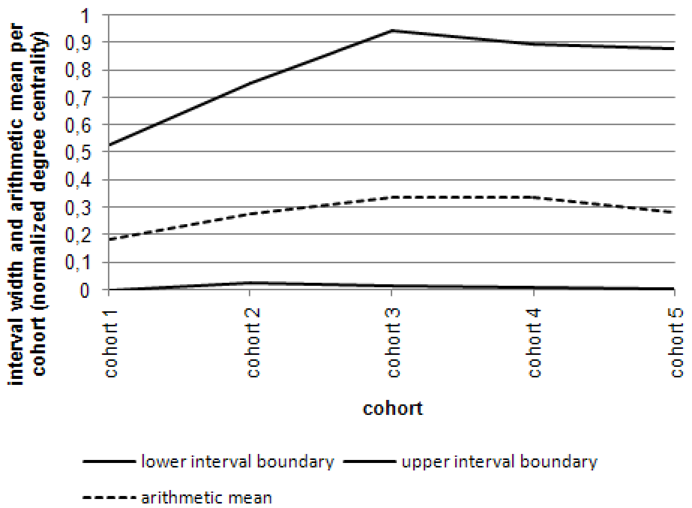
A first sign of potentially present central and peripheral actors can be deduced from the *distribution* of the normalized *degree centrality* to the classes and cohorts. The distribution reveals that – regardless of the observed cohort – the majority of companies have a normalized degree centrality ranging between 0.000 and 0.400. Some have a medium normalized degree centrality ranging between (0.400; 0.600], but few companies have a normalized degree centrality of larger than 0.600 or even 0.800. The few companies with a higher normalized degree centrality are the most central actors in the cohorts while the larger number of companies with a low normalized degree centrality can be regarded as rather peripheral actors. Particularly in cohort 3, 4 and 5 there seem to be a couple of central actors. The distribution to the classes and cohorts furthermore explains the course of the *arithmetic mean* of the normalized degree centralities: Varying between 0.183 in the first and 0.338 in the fourth cohort, the arithmetic mean remains below 0.400. The *standard deviation* furthermore shows the extent to

117 Multiple lines and loops are removed.
118 Arithmetic mean=density (see p.88, Table 25).
119 Standard deviation.
120 Degree centralization.

which the normalized degree centralities deviate from the arithmetic mean. As the computations show, it is more or less the same in all cohorts.

A more profound picture of centrality can be drawn from the *intervals*, in which the normalized degree centrality of the actors ranges. On the basis of Table 26 (see p. 97), the course of the lower and upper interval boundary as well as the arithmetic mean are shown in Figure 25 (below).

Figure 25: Interval width and arithmetic mean in all cohorts, normalized degree centrality.



Source: Own *computations based on dataset 2.*

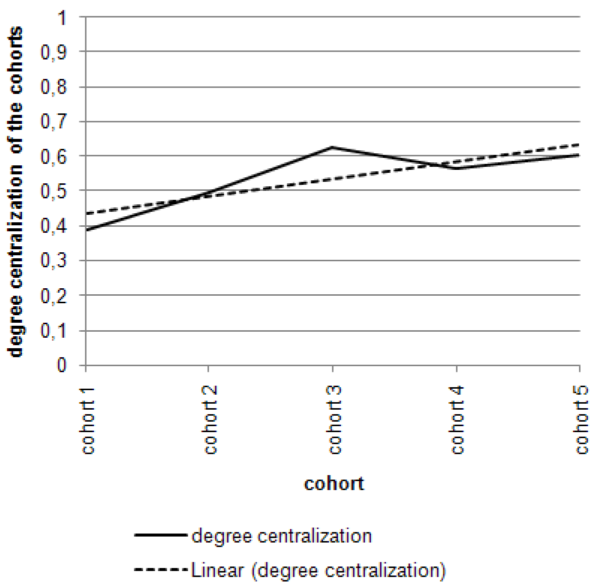
When examining the lower boundaries of the intervals it is striking that the lower boundaries are close to zero in all cohorts – ranging between 0.000 in the first and 0.025 in the second cohort. This implies that in all cohorts there are companies either having no or only few technologies with others in common. The upper boundary of the intervals shows that especially in the first cohort, there is no company with a normalized degree centrality close to one. Obviously, in these times the companies pursue rather different technologies; they are technologically not too strongly (directly) linked to each other. For instance, in the first cohort, the company with the greatest technological overlap has (at least) nine technologies with nine (of 17) other companies in common. A “break” becomes visible after the second cohort: As networks become larger and the number of patent applications increases (see p. 75, Figure 17), the upper boundary of the normalized degree centrality is constantly above 0.850. In the third, fourth and fifth cohort, there are at least three companies with a high technological overlap to other companies. E.g. in the third cohort, one company has at least 67 technologies with 67 (of 71) other companies in common.

The course of the lower and upper boundary of the intervals shows that in the first two cohorts, the gap between companies with a low and those with a high normalized degree centrality is comparably smaller than in the remaining three cohorts. Until cohort 3, deduced from the normalized degree centrality of the actors, networks seem to become increasingly centralized. Amongst all networks, especially the network of cohort 3 seems to be the most centralized one.

The analysis of the corporate landscape shows that – according to the measure normalized degree centrality – especially large chemical companies seem to continuously expand their technological connectivity to other companies. They seem to become "technological diversifiers" in the field of nanotechnology. For example, one company shows the following normalized degree centralities: cohort 1. 0.529, 2. 0.675, 3. 0.845, 4. 0.895, 5. 0.878. In three of five cohorts, this company determines the upper boundary of the respective interval; it is the most central actor in the network of technological overlap – at least according to the degree-based measurement concept. In cohort 2 and in cohort 3, two other companies determine the upper boundary.

Finally, the normalized *degree centrality values* of the companies can be held against the *degree centralization* of the entire network to gain an overall impression of network centrality. For reasons of illustration, in addition to Table 26 (see p. 97), the degree centralization of the networks is shown in Figure 26 (below).

Figure 26: Development of degree centralization.



Source: *Own computations based on dataset 2.*

Hence, the actual course of the *"degree centralization"* and the resulting trend line (dashed line) confirm the perception of increasingly centralized networks over time. However, compared to the theoretical maximum possible degree centralization of one, none of the networks can be regarded as highly centralized. Compared to a related work, namely CANTNER AND GRAF (2006), the following results can be revealed: CANTNER AND GRAF (2006: 469) who analyze the evolution of the innovator network of Jena[121] between 1995 and 2001, compute a degree centralization of 0.602 for the network of technological overlap in the time span of 1995-1997 and a degree centralization of 0.717 in the time span of 1999-2001. On average, this yields a degree centralization of approximately 0.660. In this book, in the comparable time span of 1996-2000 (cohort 4), with a value of 0.566, the degree centralization is a little lower (see p. 97, Table 26), meaning in case of nanotechnology, actors are more decentrally organized.

3.5.2.2 Implications of the centrality/ centralization measures

Previous findings suggest that especially in the last three cohorts, there are few companies which prove to be a little more central to the network of technological overlap than others: These companies have higher normalized degree centrality values than the remaining actors or, differently put, have the greatest *direct* technological overlap to other companies (which is an indication for their high technological diversity).[122]

The position of the detected "central" actors may come along with diverse implications. First of all, companies in such a position seem safer from failing (at least in the area of nanotechnology): On the one hand the application of several technologies may secure survival because even in case of the unsuccessful pursuit of one technology there are always alternatives to follow up. On the other hand, by ingeniously "combining" various technologies, new options for inventions may arise. The resulting large pool of generated technological know-how may help to foresee and influence the future of nanotechnology. In such a rapidly changing environment this seems particularly useful. All reasons perhaps account for the persistency of the central actors in the networks and also explain their strong growth[123]. Yet, a systematic investigation of the relation between technological diversity and technological survival of companies has not taken place. This, however, is subject to investigation in section 3.5.3.

Second, the technological scope and inventiveness of the central actors may also provide an ideal seedbed for (future) cooperations: Under the assumption of each company's technological focus (in principle) being known to others, it is possible

[121] See section 3.2 for further remarks.
[122] Since both, the closeness- and betweenness-based measures, are difficult to interpret in terms of the networks of technological overlap, the following conclusions are mainly drawn from the degree-based measures. Nevertheless, the results for the closeness- and betweenness-based measures are reported in section 5.3.6.
[123] Other factors, such as the already respectable size and capital of the companies when beginning their work in the area of nanotechnology may also contribute to their persistency and growth in the networks.

for the central actors to search for companies complementing their technological scope to for example further expand their position and to set standards. In connection with communication networks Granovetter (1974: 52-53 cited after GRANOVETTER 1983: 205) says:

> ""[…] Those to whom one is closest are likely to have the greatest overlap in contact with those one already knows, so that the information to which they are privy is likely to be much the same as that which one already has" […]."

Transferred to the network of technological overlap, technologically close (or in the extreme case *identical*) firms may have access to the same pool of information or resources as the central actor. Resulting, they may not contribute to the enlargement of the central actors' knowledge pool and respectively, be not valuable for its technological evolution. Accordingly, the central actor could look for companies which are active in technological fields other than its own in order to evolve and eventually, to survive. It may be helpful though if there is a common technological foundation to build upon. When two companies are technologically not related at all, finding a common basis might be difficult.

In turn, the presence of rich know-how could make the central actors attractive to other firms, especially to entrants to the field of nanotechnology. Entrants might seek the proximity to already established firms because, when entering cooperation, sale is rather guaranteed. On the other hand, cooperation between companies of significantly different sizes can lead to the absorption of the smaller partner, e.g. in the sense of acquisition.

While inventiveness and the resulting technological evolution can be regarded as desirable, with growing importance of large chemical companies the evolution of nanotechnology seems to become increasingly influenced by distinct companies or rather *the branch of chemistry*. This could not only lead to a lower diversity of inventions but in the extreme case also lead to a monopoly situation: The high exit dynamic (see p. 88, Table 25) *could* imply that some companies were simply forced out of the "market" as they were not able to successfully establish their inventions.

Whereas the normalized degree centrality values seem to point to an increasing dominance of distinct companies, the degree centralization relativizes this impression to a certain extent: None of the networks is highly centralized. Altogether, hypothesis 5 can be considered as partially true: Core and peripheral actors may be detected in the cohorts, but this pertains rather to the last three cohorts. However, despite an increasing tendency of centralization, all networks are rather far away from being highly centralized.

3.5.3 Survival in the network of technological overlap

Having discovered a slight tendency towards a center and a periphery in the cohorts, the cohorts are now examined for their survival patterns. The hypothesis is:

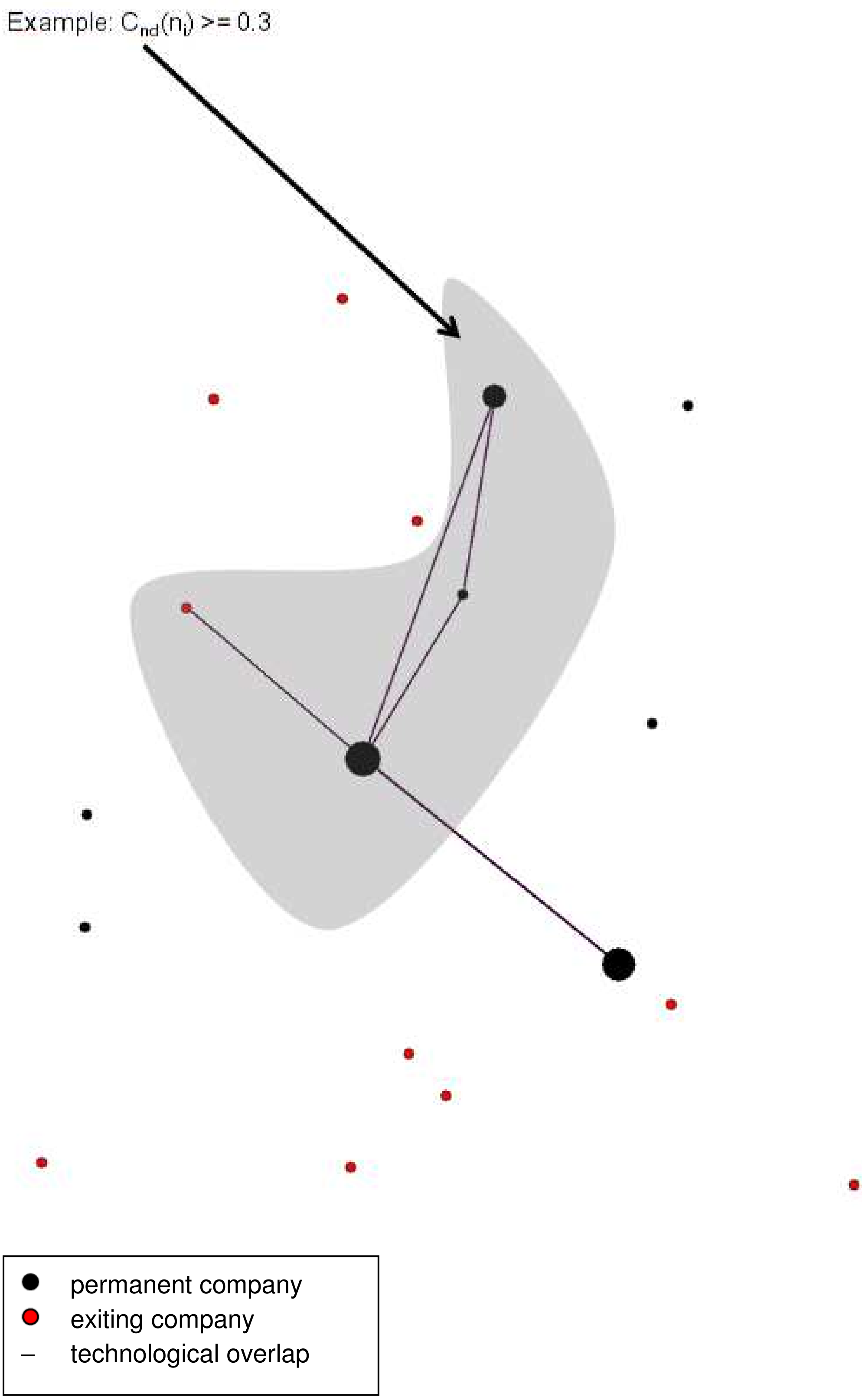
HYPOTHESIS 6: The majority of companies in the core of the network in cohort t remain actors of the network in cohort (t+1). The majority of companies in the periphery of the network of cohort t exit after cohort t, i.e. are not part of the network in cohort (t+1).

By definition, the analysis of this hypothesis is restricted to cohorts 1 to 4, because in cohort 5, the reference network of cohort (t+1) is missing. Thus, permanent and exiting actors cannot be determined in cohort 5 (see p. 88, Table 25). As mentioned in section 3.4.3, there are two ways of testing hypothesis 6: It can be tested *graphically* by contrasting the *position* (core/ periphery) against the *status* (permanent/ exiting) of companies in the networks and it can be tested *analytically* by making use of the normalized degree centralities of the actors. In a supplementary investigation, in a third step, the technological focus areas of core and peripheral companies can be contrasted against each other so that possibly, additional insight into reasons of survival can be gained.

(1) Graphical assessment

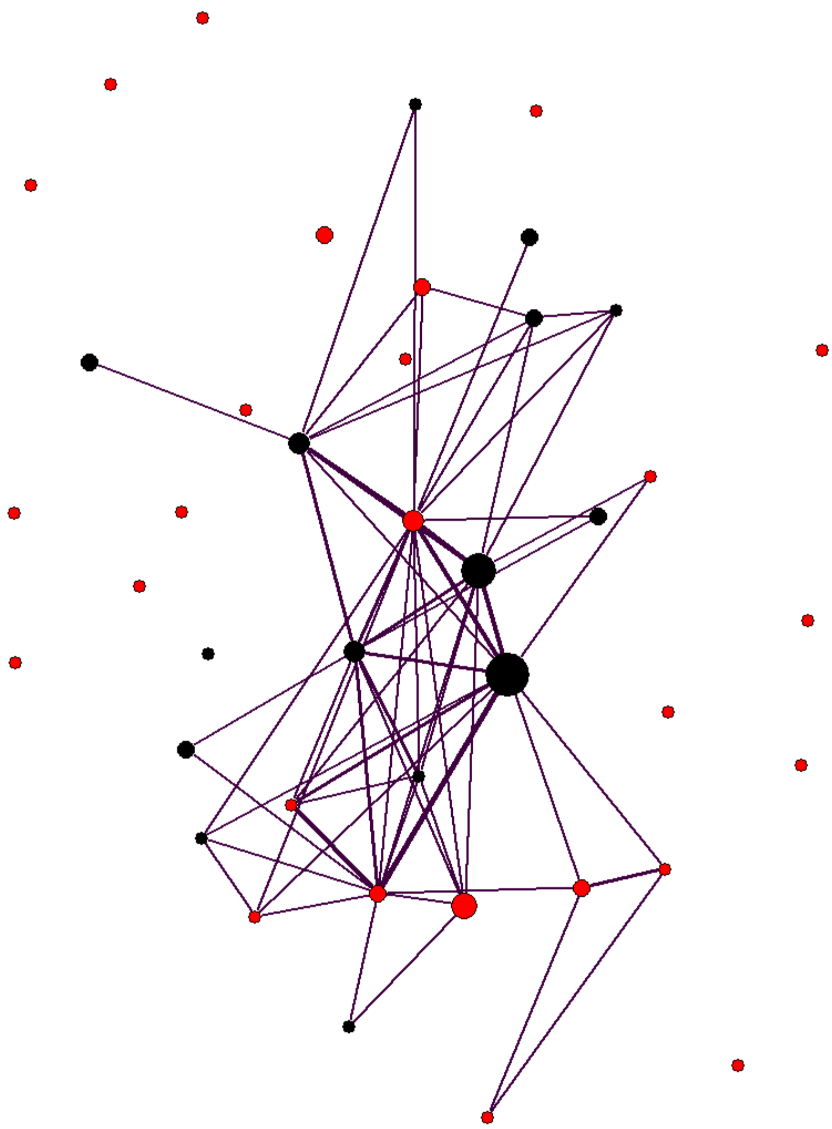
The results of the graphical assessment are presented in the following Figure 27 to Figure 30. In the figures, black vertices depict permanent actors and red vertices depict exiting actors. Else, the figures can be read as Figure 20 to Figure 23 (p. 91-94). In Figure 27, some actors are drawn against a grey background. All of these actors have a normalized degree centrality of 0.3 or higher. Concretely, this differentiation between actors relates to an example which is explained and referred to in later paragraphs.

Figure 27: Permanent and exiting companies, cohort 1.



Source: Own computations based on dataset 2.

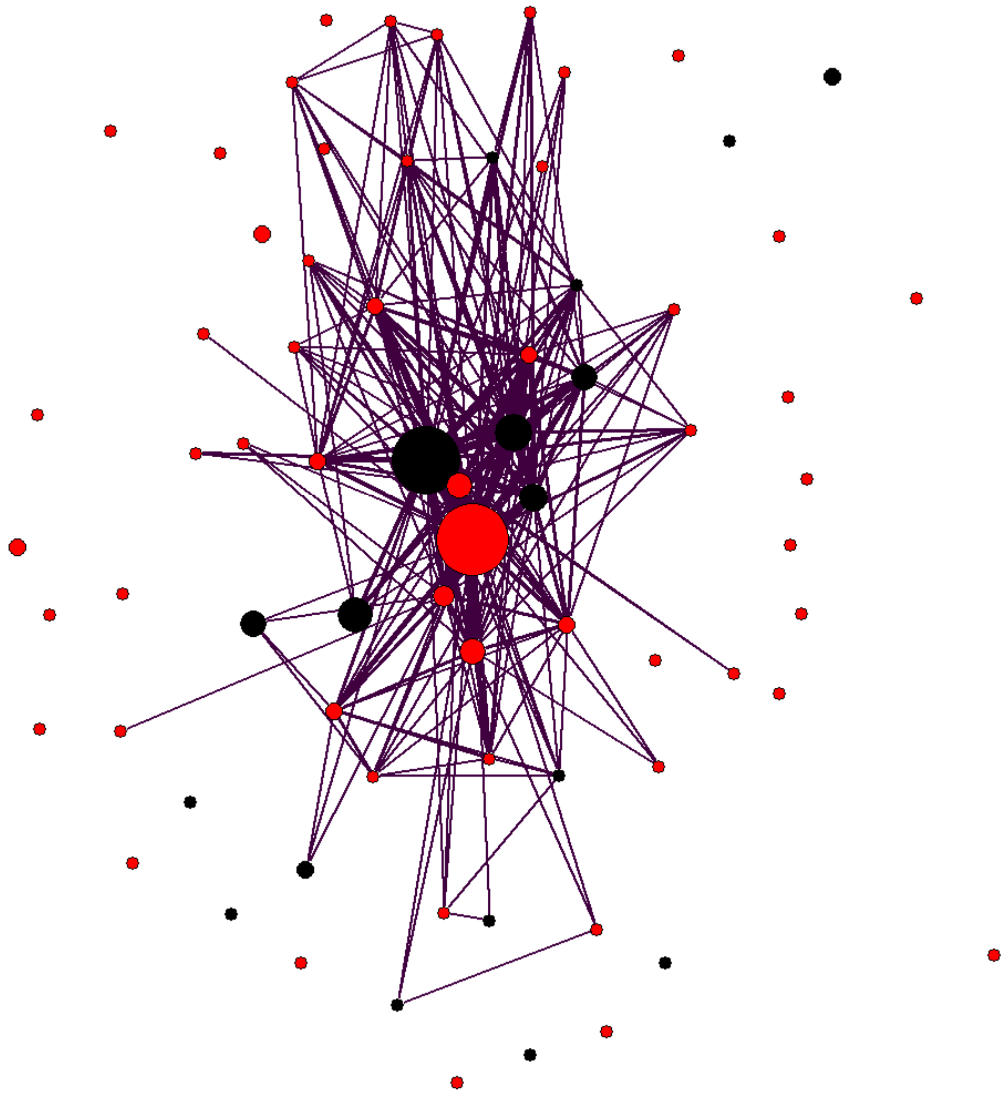
Figure 28: Permanent and exiting companies, cohort 2.



●	permanent company
●	exiting company
–	technological overlap

Source: *Own computations based on dataset 2.*

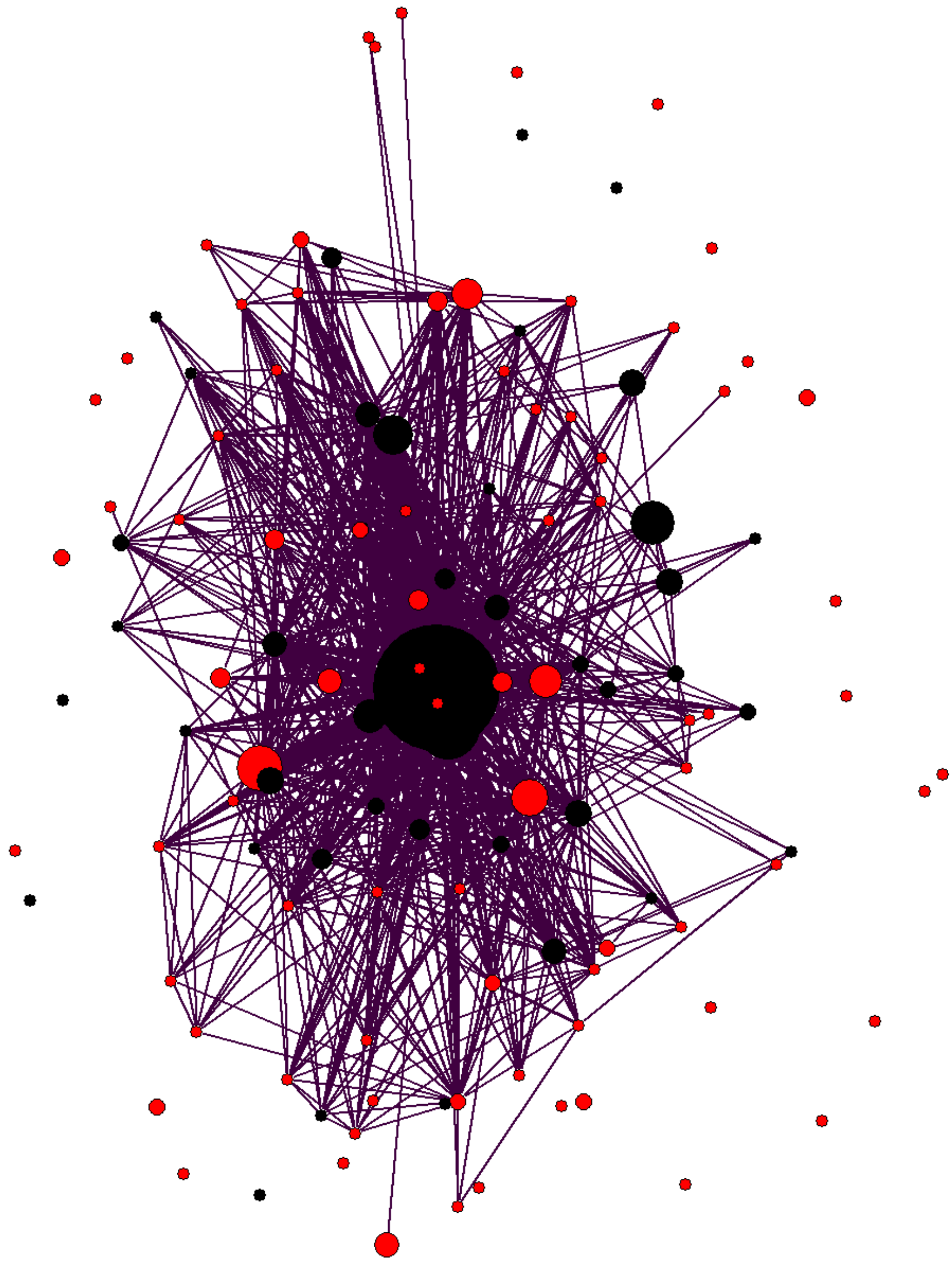
Figure 29: Permanent and exiting companies, cohort 3.



●	permanent company
●	exiting company
–	technological overlap

Source: Own computations based on dataset 2.

Figure 30: Permanent and exiting companies, cohort 4.



●	permanent company
●	exiting company
–	technological overlap

Source: O*wn computations based on dataset 2.*

The networks reveal that companies in the core (or center) of the networks tend to be permanent actors (black vertices), i.e. are also actors in the subsequent cohort. Exiting actors (red vertices), i.e. actors which are not present in the subsequent cohort, are mostly in the periphery of the networks. However, in all networks, there are also core actors which exit after the cohort or peripheral companies which remain actors in the subsequent cohort. An examination involving a thorough and systematic distinction between core and peripheral actors needs to be performed.

(2) Analytical assessment

While the graphical solution can only *subjectively* confirm or reject hypothesis 6 and furthermore, because it is not too exact in terms of defining core and peripheral actors, it is useful to consider an *analytical* approach as well. Analytically seen, such an analysis requires two kinds of information for each company to be present: First of all, it needs to be clear, whether a company is a *core* or a *peripheral* actor in a cohort. So far, actors are rather "intuitively" perceived as being "core" or as being "peripheral" actors. In this work, the assignment to the categories is guided by the normalized degree centrality of companies: Companies with a normalized degree centrality of lower than a value "x" are labeled "peripheral" companies. Respectively, companies with a normalized degree centrality equal to or higher than x are labeled "core" companies. For example, in Figure 27 (see p. 103) core and peripheral actors are distinguished employing a normalized degree centrality of 0.3. The grey area highlights the core actors – each of these companies has a normalized degree centrality of at least 0.3. The remaining actors are considered peripheral companies. They have a normalized degree centrality of less than 0.3. As shown in the following, the choice of "x" becomes a critical task. Second, information on the companies' status ("permanent" or "exiting") needs to be present. In this respect, Table 25 (see p. 88) provides a summary.

With both kinds of information available, a contingency table following Table 27 (below) can be established for all cohorts. In order to confirm or reject hypothesis 6, the percentage of *permanent* core (peripheral) actors needs to be contrasted against the percentage of *exiting* core (peripheral) actors.

Table 27: Contingency table.

	Permanent	**Exiting**
Core (normalized degree centrality >= x)	% of core companies which are permanent actors	% of core companies which are exiting actors
Peripheral (normalized degree centrality < x)	% of peripheral companies which are permanent actors	% of peripheral companies which are exiting actors

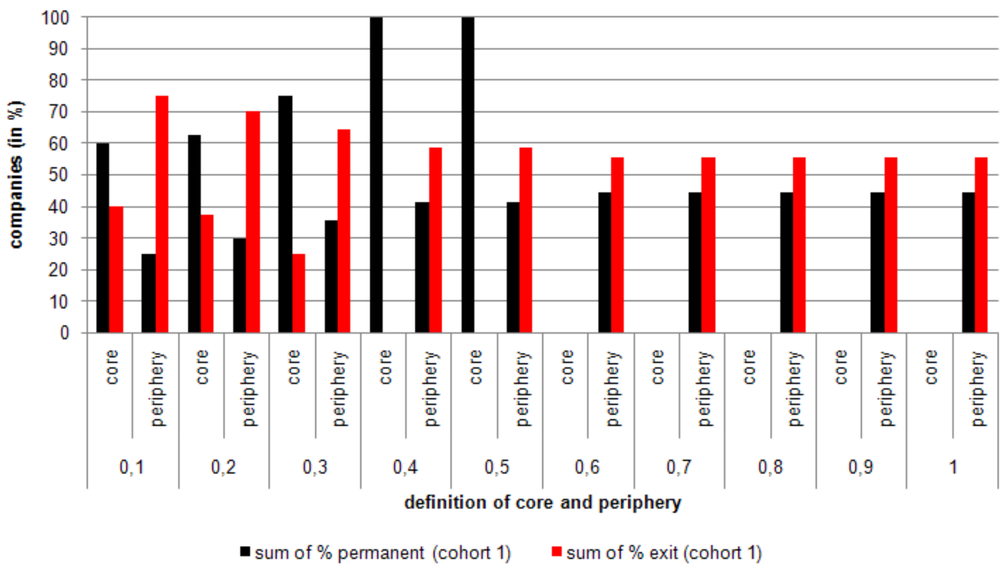
Source: Own determination.

For example, in cohort 1 employing a normalized degree centrality of 0.3 implies that three of four core actors survive while one company exits after cohort 1 (see p. 103, Figure 27). The opposite picture is revealed for the periphery: five of fourteen peripheral actors remain actors in the subsequent cohort while nine of fourteen companies exit.

The difficulty with the resulting survival pattern is that it is likely to be dependent on the "choice" of the boundary value "x": A lower boundary value implies that by tendency, more actors are defined as being "core", fewer actors are defined as being "peripheral" actors. In turn, a higher boundary value implies that, by tendency, fewer actors are defined as being "core", while more actors are defined as "peripheral" actors. Resulting, depending on the groups' composition, the share of permanent and exiting companies may vary significantly with the choice of the boundary value. Concretely, it is possible for hypothesis 6 to be confirmed for one boundary value while being rejected for another. In order to yield information on how "sensitive" the results are, a sensitivity analysis is useful. It may reveal, in how far the observed survival patterns are dependent on the choice of the boundary value.

Figure 31 to Figure 34 (below) show the results of the sensitivity analysis for cohorts 1 to 4. In the figures, the x-axis distinguishes amongst boundary values. For each boundary value, core and peripheral actors are presented separately. For both – core as well as peripheral actors – the percentage of permanent actors (black bars) and the percentage of exiting actors (red bars) is given. The percentage instead of absolute numbers is given in order to neutralize the effect of different network sizes.

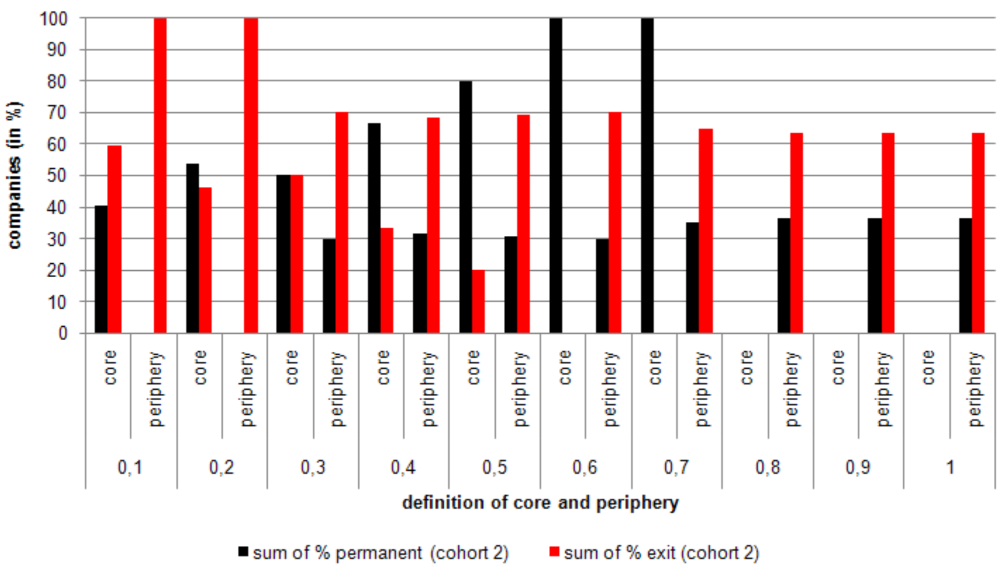
Figure 31: Sensitivity analysis for cohort 1.



Source: Own computations based on dataset 2.

As can be deduced from Figure 31, in terms of cohort 1, for example, a boundary value of 0.3 (a boundary value of 0.3 means that a company has to be connected to at least 30% of the other companies in the network in order to be declared a core company) implies that 75.000%% of the core actors remain actors in the second cohort while 25.000% of the core actors exit after the first cohort. At the same time, 35.714% of the peripheral actors remain actors in the subsequent cohort while 64.286% exit after the first cohort. This precisely reflects the picture drawn by Figure 27 (see p. 103) where three of four core actors survive while one does not and five of fourteen peripheral actors remain actors in the subsequent cohort while nine of fourteen companies exit. For the boundary value of 0.3, hypothesis 6 is therefore confirmed. As can be deduced further, hypothesis 6 is also confirmed for the remaining boundary values. If the boundary value is set to 0.6 or higher though, no company is declared a core company; all companies have a normalized degree centrality of below 0.6 (see p. 97, Table 26). Accordingly, they are all labeled as being labeled as "peripheral" companies. Since in cohort 1 (but also in all other cohorts), the number of exiting actors exceeds the number of permanent actors (see p. 88, Table 25), for these boundary values, the respective red bars are always a little higher than the black bars. This should be kept in mind when interpreting the results. Figure 32 (below) shows the results of cohort 2.

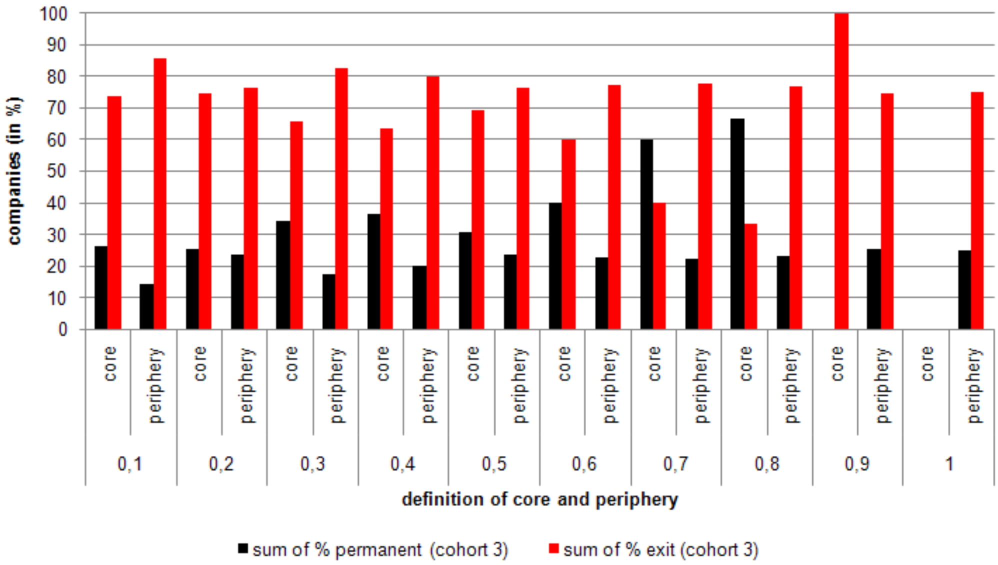
Figure 32: Sensitivity analysis for cohort 2.



Source: Own computations based on dataset 2.

For cohort 2, a similar picture is revealed, albeit this is not as clear cut as the picture of cohort 1. As can be deduced from Figure 32, hypothesis 6 is confirmed for the boundary values 0.2, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0. For the boundary values 0.1 and 0.3, hypothesis 6 is only partially confirmed. Similar to cohort 1, companies with a normalized degree centrality of higher than 0.8 do not exist in cohort 2, so for these boundary values, the bars for core actors are equal to zero. As in case of cohort 1, also in case of cohort 2, the number of companies exiting from cohort 2 exceeds the number of permanent actors (see p. 88, Table 25), so again, the results for these boundary values should be treated with care. Now, the sensitivity analysis is performed for cohort 3. The results are shown in Figure 33 (below).

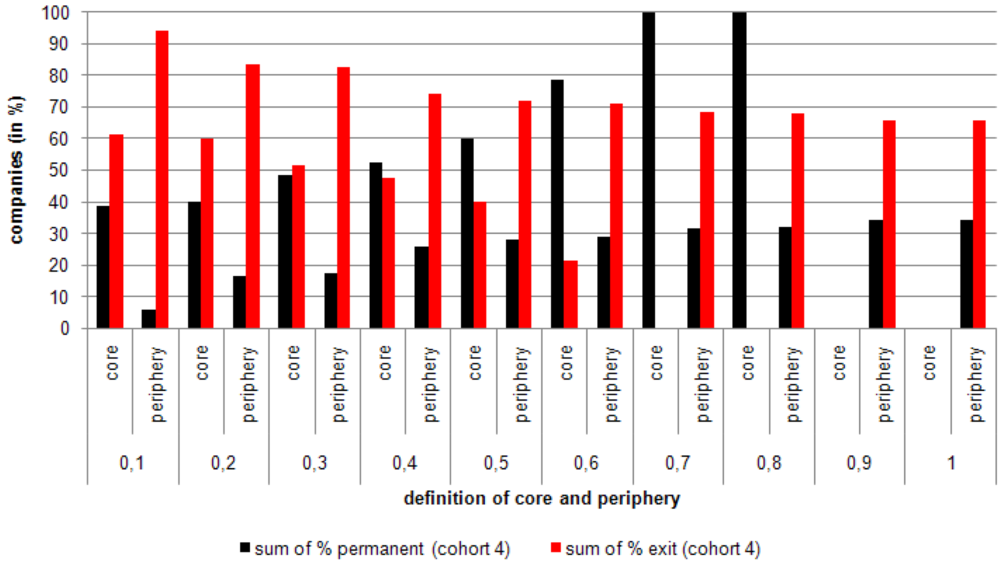
Figure 33: Sensitivity analysis for cohort 3.



Source: *Own computations based on dataset 2.*

In case of cohort 3, hypothesis 6 is confirmed for the boundary value of 0.7, 0.8 and 1.0. For the remaining boundary values, hypothesis 6 is partially confirmed. As in case of the previous cohorts, the number of companies exiting from cohort 3 exceeds the number of companies which remain active (see p. 88, Table 25). Lastly, the survival patterns of cohort 4 are under investigation. The respective results are shown by Figure 34 (below).

Figure 34: Sensitivity analysis for cohort 4.



Source: *Own computations based on dataset 2.*

Finally, the sensitivity analysis for cohort 4 shows that hypothesis 6 is confirmed for boundary values of 0.4 and higher. For boundary values of below 0.4, hypothesis 6 is partially confirmed. As in the other cohorts, also in case of cohort 4, the number of exiting companies exceeds the number of permanent companies (see p. 88, Table 25), so at least the results for the boundary values 0.9 and 1.0 should be treated with care.

Altogether, the choice of the boundary value indeed influences the resulting survival patterns. Separately seen – hypothesis 6 is always clearly confirmed for some boundary values. In these cases, technological diversity seems to be connected to higher survival chances while technological "specialization" or "isolation" rather leads to exit. For all other boundary values, hypothesis 6 is partially confirmed. When the boundary value is set to 0.8 (1.0), hypothesis 6 is confirmed in all four cohorts. This impression needs to be relativized though, as especially in cohort 1 and cohort 2, no company has a normalized degree centrality of 0.8 or higher. Consequently, all companies are labeled as being peripheral companies. In these cases, the height of the bars reflects the general distribution of permanent and exiting companies in the cohorts. Since in all cohorts, the number of companies exiting from the cohort exceeds the number of companies which remain active in the subsequent cohort (see p. 88, Table 25), the percentage of exiting companies is always higher.

That *indeed* there must be a difference between permanent and exiting actors can further be illustrated by the mean normalized degree centrality of permanent and exiting actors (see Table 28 below).

Table 28: Mean normalized degree centrality of permanent and exiting companies, cohorts 1-4.

	Permanent	Exit
Cohort 1	0.243	0.135
Cohort 2	0.385	0.214
Cohort 3	0.393	0.319
Cohort 4	0.446	0.282

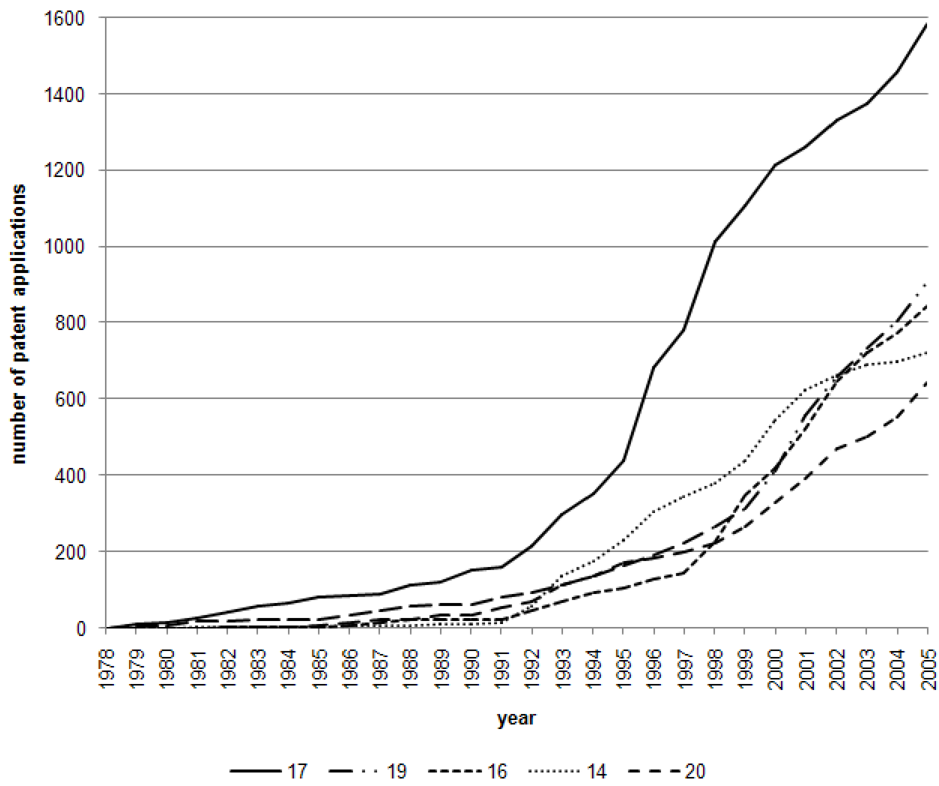
Source: *Own computations based on dataset 2.*

The analysis of ties between permanent actors and respectively, exiting actors of the networks shows that – measured by the normalized degree centrality – for all four cohorts, the ties within the group of permanent actors are considerably stronger than within the group of exiting companies. This strengthens the impression that having many technological overlaps to others protects from exit at least to some extent. Based on the findings above, hypothesis 6 can be regarded as (partially) confirmed.

(3) Supplementary investigation: Technology-related survival in the networks of technological overlap

Aside from examining exit patterns and analyzing ties, another question which might emerge is whether – aside from the mere finding of the *number* of overlaps being of importance – maybe the *kind* of chosen technology is of relevance for the technological survival of companies. It is striking that the diverse technological fields are referred to to a very different extent. Figure 35 (below) shows the development of patent applications in the top five chosen technological fields.

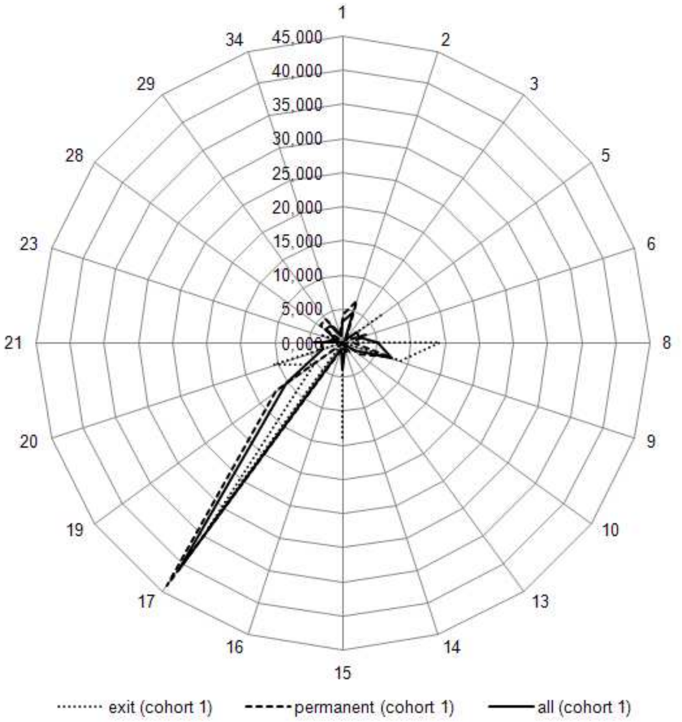
Figure 35: Development of top five chosen technologies.



Source: Own computations based on dataset 2.

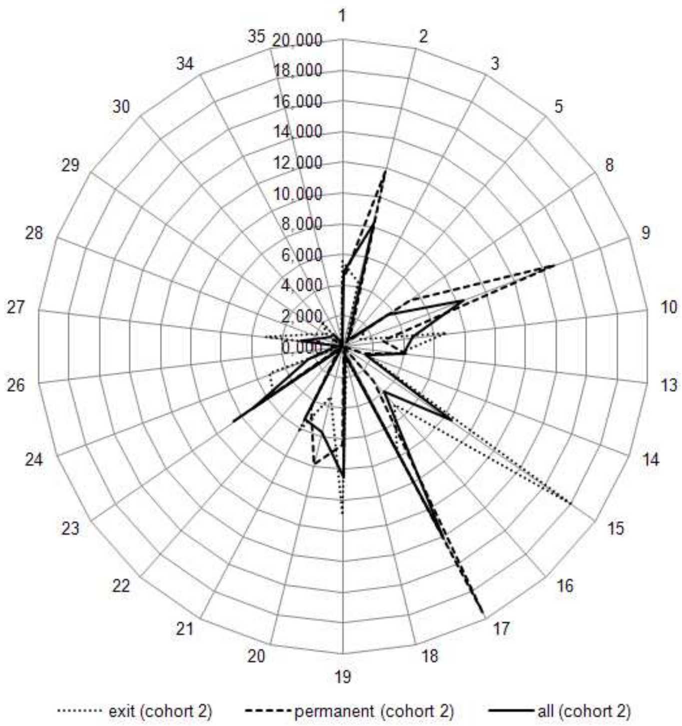
Other fields, such as field 4 (digital communication) are only sparsely referred to in the database and experience a rather moderate growth. For example, field 4 is only referred to once in the database. The different development of technological fields drives the question of whether permanent actors maybe have other focal points than exiting companies securing their technological survival. This makes it particularly interesting to contrast the technological profile of permanent actors against the technological profile of exiting companies. Figure 36 to Figure 39 (below) show the technological profile of both.

Figure 36: Technological profile of cohort 1, in percent.



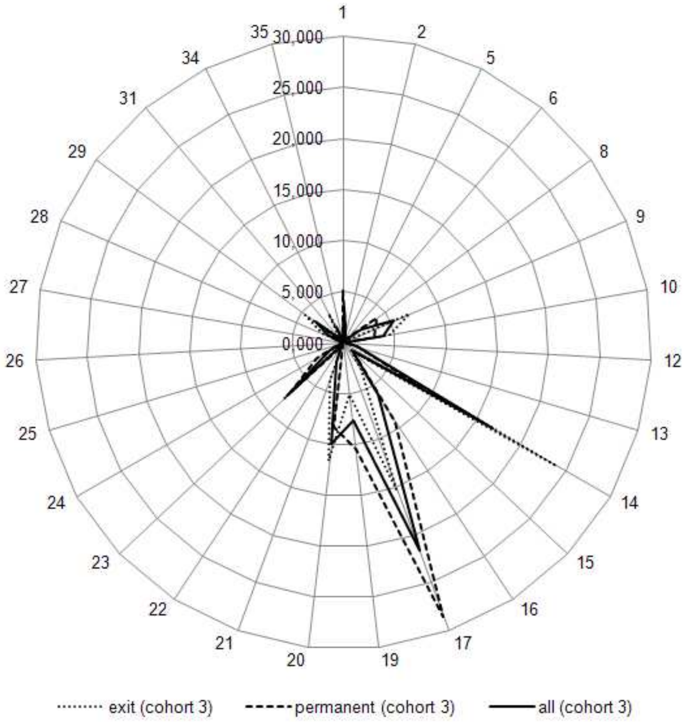
Source: Own computations based on dataset 2.

Figure 37: Technological profile of cohort 2, in percent.



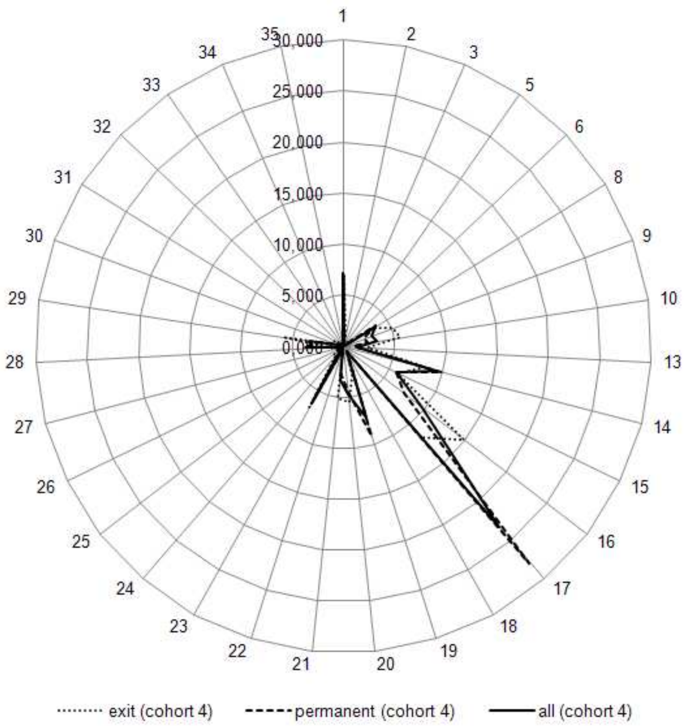
Source: Own computations based on dataset 2.

Figure 38: Technological profile of cohort 3, in percent.



Source: Own computations based on dataset 2.

Figure 39: Technological profile of cohort 4, in percent.



Source: Own computations based on dataset 2.

Counting the numbers (each representing a technological field, see p. 77, Figure 18) in the outer circle of the figures shows that, in general, the number of chosen technologies grows: In cohort 1, altogether, 20 technologies are chosen, in cohort 4, 31 technologies are being referred to. In accordance with Figure 18 (see p. 77), some fields stand out: Mostly, this is field 17 (macromolecular chemistry, polymers) which refers to the chemical aspects of polymers (see SCHMOCH 2008: 13). However, also other technological fields, such as 14 (organic fine chemistry) and 15 (biotechnology) are being referred to to a greater extent.

In all four observed cohorts, permanent and exiting actors have a *similar* but not *identical* technological profile: Interestingly, in all cohorts, permanent actors always devote the majority of their efforts to field 17 (macromolecular chemistry, polymers). In the first cohort, exiting companies also have their main emphasis on field 17. In the second cohort they concentrate on field 15 (biotechnology). In the third cohort exiting companies focus on field 14 (organic fine chemistry) and in the fourth cohort, they mostly refer to field 16 (pharmaceuticals).

The second emphasis of permanent and exiting companies differs across the cohorts. For instance, in the first cohort, permanent actors focus on field 19 (basic materials chemistry) second most, while exiting companies concentrate on field 8 (semiconductors) and 15 (biotechnology) second most. In cohort 2, permanent actors put their efforts into field 9 (optics) while exiting companies choose field 19 (basic materials chemistry) which was chosen to be the second emphasis of permanent companies in the first cohort. In the third cohort, permanent companies again put effort into field 19 (basic materials chemistry), while exiting actors focus on field 17 (macromolecular chemistry, polymers). Finally, in cohort 4, permanent actors also put their efforts into field 14 (organic fine chemistry) while exiting companies remain focusing on field 17 (macromolecular chemistry, polymers).

In conclusion, permanent and exiting actors do have a slightly, but not very systematically varying technological profile. Next to the technological diversity of companies, also the choice of distinct technologies might therefore partially account for the persistency of some companies within the networks or respectively, the exit of others. However, both “invest” into strongly growing technologies. The underlying assumption of exiting companies being involved in lesser chosen technologies cannot be confirmed. Since “exit” in this main section is defined as “technological exit” and not in the sense of failure, it is possible for the respective companies to have shifted their focal points to areas outside of nanotechnology.

3.6 Summary and conclusion

Further aiming at explaining the evolution of domestic nanotechnology companies, in section 3, an *alternative* perception of survival, concretely, a *technology-driven* approach is chosen: It is analyzed whether the general choice of a company’s technological orientation (depicted on patent specifications) may account for a company's *technological survival*. Other than in section 2, entries and exits are defined over the occurrence of patent applications which allow for repeated entries into and exits from the network of technological overlap. Concretely, three hypotheses are under investigation:

> ***HYPOTHESIS 4:*** The structure of the network of technological overlap changes strongly over the observed cohorts in terms of the actors which are part of the network and in terms of the intensity of ties between the actors.
>
> ***HYPOTHESIS 5:*** In each period of time, in the networks, there are few diversified actors which – by means of social network analysis – are clearly identifiable as core actors with a high technological overlap to other actors. Consequently, there are also companies in the periphery of the network with a small technological overlap to other companies.
>
> ***HYPOTHESIS 6:*** The majority of companies in the core of the network in cohort t remain actors of the network in cohort (t+1). The majority of companies in the periphery of the network of cohort t exit after cohort t, i.e. are not part of the network in cohort (t+1).

When it comes down to such a closer investigation of the technological evolution, profound information on the companies and their technological scope needs to be present. Since the database used in section 2 does not allow for such an in-depths examination, in this section, data from the database PATSTAT (VERSION 10/ 2007) is used, referred to as dataset 2. Extracted using a comprehensive search strategy following Fraunhofer, ISI (in: NOYONS ET AL. 2003: 100-101, see section 5.2.2), the data is prepared such that altogether 1284 patent applications (with a priority date between 1978 and 2005) applied to at the EPO/ WIPO are retrieved deriving from 382 domestic companies. The patent applications are assigned to 35 technological fields following the classification scheme as suggested by SCHMOCH (2008: 9-10). The data is furthermore split into five cohorts.

Inspired by CANTNER AND GRAF 2006, to test the above hypotheses, a social network analysis is performed consisting of three steps: In preparation of answering hypothesis 4, in a first step, based on the retrieved nanotechnology patent applications, a network of technological overlap is constructed for each of the five cohorts (as described in section 3.2 in more detail, in the networks of technological overlap, the actors are companies and ties between the companies depict a present technological overlap between them). Basic structural attributes such as the number of actors and their division into permanent, exiting and entering actors are considered. Also the ties between actors are analyzed. The second part of the analysis is concerned with the detection of centrality in the networks of technological overlap and its possible implications. This step is essential for deriving a statement concerning hypothesis 5. Third, methods for detecting survival patterns in the networks of technological overlap are used to correspond to hypothesis 6. Altogether, the results – organized in the order of the hypothesis – are summarized in the following.

In order to derive a statement concerning hypothesis 4, diverse descriptive statistics are computed from the five networks of technological overlap. They show that in the course of time the networks of technological overlap experience a remarkable change which can also be verified by the visual impression. While – except for the third cohort – the company holding the majority of patent applications remains the same over all cohorts and the following two ranks are usually filled with mostly large chemical companies, in general, the strongly growing numbers of actors and the constantly high entry and exit rates indicate that else, the networks underlie high dynamics. Concerning the ties between the actors, the average degree and density depict that the companies tend to be increasingly connected to each other. Hypothesis 4 is therefore confirmed.

Concerning hypothesis 5 the examination shows, that especially in cohorts 3, 4 and 5 there are at least three companies which – according to the measure normalized degree centrality – are a little more central in the networks of technological overlap than the remaining companies. According to the normalized degree centrality, the top three companies – generally, large chemical companies – are mostly alike in all cohorts. This is not surprising as in all cohorts, these companies usually are amongst the companies with the highest number of patent applications: With many patent applications, the probability of covering many technological fields is often higher. In general though, none of the depicted networks is technologically highly centralized. However, the trend line of the degree centralization depicts a growing tendency towards centralized networks over time. Altogether, hypothesis 5 can be regarded as partially true: Core and peripheral actors are detectable in at least the later three cohorts. Nevertheless, again, all networks are far away from being highly centralized.

Finally, the survival patterns of core and peripheral actors are under investigation to derive a statement in terms of hypothesis 6. Due to a missing reference network, permanent and exiting actors are not determinable in cohort 5, so hypothesis 6 can only be investigated for cohorts 1 to 4. The visual impression reveals that in cohorts 1 to 4, companies in the *core* of the networks rather remain active in the subsequent cohort while companies in the *periphery* of the networks tend to be more likely to experience a technological exit.

Since the graphical solution rather allows for a subjective assessment of the survival patterns, and furthermore, core and peripheral companies are not distinguished in a systematic manner, an *analytical* approach is then performed. A sensitivity analysis is carried out in order to show in how far the *choice* of the *boundary normalized degree centrality* between core and peripheral actors is likely to influence the resulting survival patterns. The results show that the choice of the boundary normalized degree centrality influences the resulting survival patterns to a certain extent: While in terms of cohort 1, hypothesis 6 is confirmed for *all* boundary values, in cohort 2, it is confirmed for eight of ten boundary values, for the remaining two boundary values it is partially confirmed. In case of cohort 3, hypothesis 6 is confirmed for three boundary values and partially confirmed for the rest. Finally, in cohort 4, hypothesis 6 is confirmed for boundary values of 0.4 and higher and partially confirmed for those below. When the boundary value is set to 0.8, implying that companies with a normalized degree centrality of 0.8 or higher are labeled core companies, hypothesis 6 is confirmed in all four cohorts simultaneously. As especially in cohort 1 and cohort 2, no company has a normalized degree centrality of 0.8 or higher though, and hence, all companies are labeled as being peripheral companies, this impression needs to be relativized though: In these cases, the sensitivity analysis merely depicts the general distribution of permanent and exiting companies in the cohorts. Since in all cohorts, the number of exiting companies exceeds the number of permanent actors, the percentage of exiting companies is always higher.

That indeed there is a difference between permanent and exiting actors can further be highlighted by their mean normalized degree centralities. In all cohorts, the ties within the group of permanent actors are considerably stronger than within the group of exiting companies. Being a core company therefore seems to protect from exit to some degree.

In summary, it can be concluded that the “choice” of a company’s technological fields depicted on its nanotechnology patent applications seems to indeed exert an influence on the company’s survival chances. Companies with a *higher* technological overlap rather remain active in the area of nanotechnology than those with a smaller technological overlap. However, the extent to which technological overlap occurs can depend on the general number of patent applications of a company. The more patents a company applies for, the higher its probability of covering many fields of technology. Altogether, hypothesis 6 can be regarded as (partially) true: The majority of companies in the core of one network remain actors of the network in the subsequent cohort, but the chosen boundary normalized degree centrality distinguishing between core and peripheral actors exerts a slight influence on the observed survival patterns.

Since in the database, technologies are being referred to each to a different extent, this raises the question whether – next to the general technological diversity – maybe other factors, such as the choice of *distinct* technological fields, lead to the technological survival of some companies or respectively, the technological exit of others. The technological profile of permanent and exiting companies visualizes that both kinds of companies have similar, but not identical technological profiles. This might account for the persistency of some companies within the networks and the exit of others. However, both “invest” into strongly growing technologies, so – recalling that *technological* exit and *market* exit cannot be equated – it is likely for the exiting companies to merely have shifted their interest to areas other than nanotechnology.

Due to the apparent relation between the technological diversity and technological survival of firms, there are several options for additional works. For example, for companies it might be particularly rewarding to thoroughly examine the networks of technological overlap: Especially *because* indications point to companies with a higher technological overlap to other companies to technologically survive, a company may strive to enlarge its patenting portfolio in order to cover more technological fields and thereby to enhance its (technological) survival chances in the area of nanotechnology. Together with the findings of previous researchers, for example POWELL ET AL. (1996), who regard cooperation between firms to be particularly fruitful in terms of fostering innovation, firms could therefore investigate the networks of technological overlap to find an ideal partner to cooperate with or even to acquire. So far, only a few actual cooperations (18 of 1284 patent applications include more than one company as an applicant) are determinable in the present database. However, a growing tendency of cooperations is perceivable (see section 5.3.3).

For future works, it could then be interesting to take a closer look at such interfirm collaborations, especially because – as SYDOW (1992: 54) said a couple of years ago – the organizational form of a corporate network gains popularity. Whether in the context of this development, the organizational form of vertically deeply integrated and/ or widely diversified large companies becomes obsolete and small and medium sized companies lose their traditional character by being integrated in such a network, SYDOW (1992) regards as doubtful. The author strengthens, that even if the corporate network is not necessarily *the* organizational form of the future, it is an organizational form *with* future.

For gaining further insight into interfirm collaborations, next to examining patent applications, boards of directors of different companies could be analyzed to determine whether the detected technological proximity of companies can be confirmed and explained by personal affiliations e.g. a person being a member in more than one supervisory board. In a further step, an investigation resembling the one by AHUJA (2000: 425) could be performed: The author assesses the effects of a firm's network of relations on innovation. Networks of personal relations (cooperation, scientist mobility) following CANTNER AND GRAF (2006: 466) could further be constructed to gain a deeper insight into interfirm relations. In either case, interfirm relations could be contrasted against the technological survival of firms.

In additional works, similar examinations could be conducted for other technologies to determine how the diverse technologies differ from another in terms of the width of their technological linkages. Similarly, cross-country comparisons could give some indication on the basis and development of nanotechnology in other countries. Possibly, in other countries, nanotechnology touches upon other technological fields than in Germany.

4 RÉSUMÉ

4.1 Summary of results

Nanotechnology is perceived as the technology of the future (see BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG 2004: 4), technology with a cross-sectional character or "enabling technology" (see BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG 2006B: 3, 11). At the same time, risks for mankind and the environment are often emphasized (e.g. see KÜHLING AND HORN 2007). Resulting, each nation's economic welfare is increasingly reliant on the technology and its development. Under the assumption that – next to gaining "technology specific" insights – suitable framework conditions need to be established to secure the technology's (sustainable) development and to support a region's competitiveness, this book sheds some light on a selection of factors possibly fostering or hindering the evolution of domestic nanotechnology companies. Two empirical approaches are presented to assess the role of pre-entry/ post-entry experience, technological know-how and the technological orientation of companies in terms of the influence they exert on the *actual* and *technological* evolution of the domestic corporate landscape. The focus is on domestic companies as they increase rapidly and therefore are not only assumed to play a key role in terms of the *technological* development but also to contribute largely to the nation's *competitiveness*. This book thereby complements previous works which – amongst others – concentrate on particular factors driving the technological change (HULLMANN 2001), the corporate (technological) evolution (BURR ET AL. 2009) or the evolution of companies involved in nanotechnology (HEINZE 2006).

The first empirical approach is presented in section 2 where the *actual survival* of firms is in the center of attention. To assess in how far the factors pre-entry experience, post-entry experience and technological know-how relate to actual firm survival in the time between 1978 and 2009, company information from five different sources[124] is collected and prepared in a time-consuming process using HOPPENSTEDT and PATSTAT (VERSION 10/ 2007) data. The remaining gaps are filled by a very extensive and therefore extremely time-consuming manual internet research involving sources such as LEXISNEXIS and GENIOS, but also a high number of – for example – corporate web pages, press releases etc. Methods of survival or duration analysis are then applied to dataset 1 comprising the 354 companies which remain for further analysis. Concretely, the duration analysis consists of two parts: Kaplan-Meier estimates and a (stratified) Cox regression. In the context of the Cox regression, several compositions of regression models

[124] BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG/ BUNDESMINISTERIUM FÜR WIRTSCHAFT UND TECHNOLOGIE: Förderkatalog, http://oas2.ip.kp.dlr.de/foekat/foekat/foekat, 16 January 2007; HEINZE (2006); VDI-TECHNOLOGIEZENTRUM: Nanomap, http://www.nanomap.de/, 7 August 2007; NANOTECHNOLOGIE IN HESSEN: Nanotechnologie-Unternehmen, http://www.nanoportal-hessen.de/brancheninfo/unternehmen/, 9 March 2007; NANOPRODUCTS: http://www.nanoproducts.de/, 9 March 2007.

(involving the three covariates) are tested. As the subsequent paragraphs highlight, the contribution of the duration analysis is manifold. In summary, the first three hypotheses yield the following picture:

> ***HYPOTHESIS 1:*** Nanotechnology companies have a higher survival probability if they are equipped with pre-entry experience.
> → *neither confirmed nor rejected*
>
> ***HYPOTHESIS 2:*** Nanotechnology companies have a higher survival probability if they are equipped with more post-entry experience.
> → *confirmed*
>
> ***HYPOTHESIS 3:*** Nanotechnology companies have a higher survival probability if they are equipped with technological know-how.
> → *partially confirmed*

In more detail, concerning *pre-entry experience* (central to hypothesis 1) – other than previous studies suggest (e.g. see KLEPPER 2002A: 661FF, CANTNER ET AL. 2006: 52, 57FF, THOMPSON 2003: 15, 27) – the results show that companies with pre-entry experience tend to have a higher survival probability for a limited time span only. Merely between (approximately) age 12 and age 24, companies with pre-entry experience are better off than companies without pre-entry experience. The regression analysis furthermore shows that pre-entry experience does not shape the observed hazard curve of the companies in the sample. *Hence, hypothesis 1 is neither confirmed nor rejected.*

Under the supposition that experience might not only be present at inception but also be gathered during a company's operation on the market, hypothesis 2 focuses on the relevance of *post-entry experience*. In this respect, previous studies detect that, by the time of the shakeout, earlier entrants have lower hazard rates persisting many years thereafter; also, hazard rates seem to decline with age in some industries while in other industries age does not exert an influence on the hazard (see KLEPPER AND SIMONS 1999: 36FF). HORVATH ET AL. (2000: 18) amongst other report elevated hazard rates for cohorts entering late in an industry's life cycle (findings for the U. S. automobile, beer brewing and tire industry). The results for nanotechnology seem to confirm the latter finding: companies with more post-entry experience have a higher survival probability or – differently put – "older" companies are rather resistant to failures than newcomers. By implication this means that the later the entry takes place, the higher the hazard of exit. The relevance of post-entry experience is directly underlined by the regression results: Whenever post-entry experience is considered in a model, it reaches high significant levels. *Based on the present data, hypothesis 2 is therefore confirmed.*

The third hypothesis aims at shedding some light on the importance of *technological know-how* (approximated by innovative activities). In this respect, previous studies suggest a correlation between innovative activities and the survival of companies. It is mentioned that innovative activities may even compensate for the lack of post- and pre-entry experience (see CANTNER ET AL. 2005: 1). In case of nanotechnology, a connection between innovative activities and firm survival is also present. It appears that, while being irrelevant or almost "unfavorable" at inception, companies with technological know-how are better off in the long run. Due to its time dependency, whenever technological know-how is considered in a regression model, the covariate is included as a stratification variable, so a concluding statement concerning its explanatory power in terms of the observed hazard cannot take place. However, compared to their "unstratified counterparts", the stratified models show slightly improved significance levels, so technological know-how seems to at least have some relevance in terms of explaining the actual firm survival. *Mainly based on the Kaplan-Meier estimates, hypothesis 3 is partially confirmed.*

In the second part of the empirical analysis presented in section 3, to better account for the fact that nanotechnology is a technology, the *technological survival* of firms is under investigation. To assess whether the technological overlap of companies accounts for technological firm survival in the time between 1978 and 2005, 1284 EPO/ WIPO-patent applications (deriving from 382 domestic companies) are retrieved from the EPO Worldwide Statistical Patent Database version October 2007 (PATSTAT (VERSION 10/ 2007)) using the search strategy developed by Fraunhofer, ISI (in: NOYONS ET AL. 2003: 100-101, see section 5.2.2). The retrieved dataset is referred to as dataset 2. The patent applications are assigned to 35 technological fields according to the classification scheme suggested by SCHMOCH (2008: 9-10) and are further prepared for analysis. Amongst other, the data is split into five cohorts (according to their priority year) altogether comprising the time between 1978 and 2005. Methods of social network analysis are then applied to the 382 companies of dataset 2. Concretely, the analysis is inspired by CANTNER AND GRAF (2006) and consists of three parts: First of all, networks of technological overlap are constructed for the five cohorts. In these networks, the actors are domestic companies and ties represent a technological overlap between the companies: Following CANTNER AND GRAF (2006: 466), whenever companies have a technological field in common, they are connected to each other. Resultingly, in each network, all companies sharing the same technology are connected to each other (such a connection should not be equalized with a cooperation). In connection with this first step, the basic network structure including the number of actors, their division into permanent, exiting and entering actors, and the ties between them is explored. Second, roles of actors (core or peripheral actor) are investigated. Third, the survival patterns of the companies are examined by combining the results of the first two methodological steps. In direct comparison to section 2, two major differences are exploited in section 3: First of all, in section 3, the factor "technological know-how" is no longer being treated as binary, but as multifaceted, i.e. as consisting of several technological fields. Second, in section 2, companies which exit do not re-enter. In section 3, companies are considered as entering, exiting or permanent actors depending on their patenting behavior. Using patent applications allows for the definition of repeated entries and/ or exits, so higher dynamics than in section 2 are permitted. Altogether, the social network analysis reveals notable findings which are briefly outlined in the subsequent paragraphs. In summary, hypotheses 4 to 6 yield the following picture:

> ***HYPOTHESIS 4:*** The structure of the network of technological overlap changes strongly over the observed cohorts in terms of the actors which are part of the network and in terms of the intensity of ties between the actors.
> → *confirmed*
>
> ***HYPOTHESIS 5:*** In each period of time, in the networks, there are few diversified actors which – by means of social network analysis – are clearly identifiable as core actors with a high technological overlap to other actors. Consequently, there are also companies in the periphery of the network with a small technological overlap to other companies.
> → *partially confirmed*
>
> ***HYPOTHESIS 6:*** The majority of companies in the core of the network in cohort t remain actors of the network in cohort (t+1). The majority of companies in the periphery of the network of cohort t exit after cohort t, i.e. are not part of the network in cohort (t+1).
> → *(partially) confirmed*

In detail, regarding the basic *network structure*, which is central to hypothesis 4, the results display that over time, the networks of technological overlap experience a remarkable change: This is determinable by the strongly growing number of companies involved in nanotechnology as well as by the constantly high entry and exit rates. The network of technological overlap of cohort 1 for example no longer exists in its original form in the following cohorts. Merely some actors remain active in the subsequent cohorts. The others exit from the network of technological overlap. Also ties between the actors change. *Hence, this leads to the verification of hypothesis 4.*

The *roles* of the actors are addressed in hypothesis 5. Since large (mostly) chemical companies always seem to hold the majority of patent applications, not surprisingly, these companies have comparably high normalized degree centrality values and – according to the degree-based measures – therefore seem to be at least a little more central to the overall network structure than others. Altogether, none of the networks is technologically highly centralized though. However, the trend line drawn for the measure degree centralization depicts a growing tendency towards centralized networks over time. *Altogether, hypothesis 5 is therefore partially confirmed.*

Finally, the *survival patterns* of core and peripheral actors (in the cohorts) are investigated for hypothesis 6. Due to a missing reference network (t+1), by definition, permanent and exiting companies cannot be distinguished in terms of cohort 5, so the hypothesis can only be tested for cohorts 1 to 4. For these cohorts, permanent and exiting companies are first distinguished graphically in the networks of technological overlap. The visual impression reveals that in cohorts 1 to 4, companies in the core tend to be permanent actors while exiting firms are mostly found in the periphery of the networks. In order to support and strengthen the visual impression, a sensitivity analysis is then carried out distinguishing between core and peripheral companies in a more systematic manner. It shows the extent to which the choice of the boundary normalized degree centrality between core and peripheral actors is likely to influence the resulting survival patterns. Concretely, the percentage of permanent and exiting companies is contrasted for core and for peripheral actors for diverse boundary normalized degree centrality values. The findings are that altogether, the majority of companies in the core of one network remain actors of the network in the subsequent cohort and that the majority of companies in the periphery of the network exit afterwards. However, the boundary drawn between core and peripheral actors exerts an influence on whether both parts of this statement or just one part are/ is confirmable. In summary, it can therefore be concluded that the choice of a company's technological fields – "choice" in the sense of choosing common or less common technological fields – depicted on its nanotechnology patent applications seems to influence the technological survival of companies at least to some degree. Critically, it should be remarked that companies with a higher technological overlap are usually the companies with a higher number of patent applications: The more patents a company applies for, the higher is its probability of covering many fields of technology. *Altogether, hypothesis 6 is regarded as (partially) confirmed.*

Next to the technological diversity of companies being relevant to explain their technological survival, it is analyzed whether permanent actors focus on other, possibly more prosperous, technological fields than exiting actors. It can be shown that permanent and exiting companies have similar, but not identical technological profiles. Both "invest" into strongly growing technologies though. The underlying assumption of exiting companies being involved in lesser relevant technologies can therefore not be confirmed. Since in section 3, "exit" is defined as "technological exit" and not as market failure, it is possible for the respective companies to have shifted their focal points to areas outside of nanotechnology.

In a nutshell, pre-entry/ post-entry experience, technological know-how and the technological orientation depicted on a company's nanotechnology patent applications seem to explain the actual and technological survival of companies to a certain degree. Altogether, the domestic corporate nanotechnology landscape seems to underlie high dynamics. Within this dynamic environment, a notion of emerging and persistent core companies is detectable, covering a wide spectrum of technological fields. Together with the finding of technological know-how being a possible success factor in the long run, framework conditions in the area of nanotechnology (at least in the present state) seem to be particularly in favor of technologically diversified companies. Especially large chemical companies seem to have emerged to be the companies sticking out in and dominating the corporate landscape (this also explains the elevated number of patent applications in the area of chemistry). They appear to shape the technological evolution of nanotechnology to a large extent. They furthermore prove to be rather old and large in size. Their age could thereby be regarded as a possible additional success factor. Based on the present state and the findings above, political recommendations and an outlook as given in the following section can be formulated.

4.2 Political recommendations and outlook

Nanotechnology is a very young technology. This comes along with the implication that long term, and particularly, more precise data material is simply not available yet. Together with the fact that nanotechnology is in its growth stage with a high degree of dynamics being present (and to be expected further), political recommendations can and should only be of tentative nature. Future works should continuously track, improve and enrich the existing data in order to confirm the findings and to observe long or at least longer term trends. Being aware of the present limitations and under the assumption that the findings of this book are representative of the entire corporate landscape, the following preliminary recommendations for political actions can be deduced:

- Since for newly founded firms it does not seem to be important to have pre-entry experience (prior experience seems only helpful in the mid-term perspective) in order to survive, at least in the present stage, there is no particular reason to set up programs to specifically foster its emergence. Potential founders should rather be motivated to get involved into nanotechnology – they should not fear to fail simply because others have prior experience in fields other than nanotechnology.

- As it appears that earlier entry is immediately connected to gathering more market know-how which helps to survive in the long run, entry should be encouraged to happen as *early* as possible. Therefore, when aiming to secure long term (domestic) firm survival in the area of nanotechnology, framework conditions should be set such that they ease foundries in the area of nanotechnology.

- Third, companies should be animated to set up a solid fundament of technological know-how. This is essential in two respects at the same time: Not only seems the mere presence of technological know-how lead to higher *actual firm survival* in the long run. If patent activities are sufficiently wide spread, so that many technological fields are touched upon (which are also covered by competitors), this also seems to secure *technological persistency* in the area of nanotechnology. Therefore, framework conditions should encourage companies to choose a not too narrow scope: At best, patent applications should touch upon diverse technological fields, so that they possibly enable and trigger manifold starting points for others to involve in nanotechnology. By ingeniously positioning the technological landscape, a seedbed for innovation can eventually be established.

Next to performing the same kind of analysis on an enriched sample, several extensions to the survival and the social network analysis are imaginable. While the examined variables deliver a first valuable insight on possible reasons of survival, for giving an even better prediction on the evolution of domestic nanotechnology companies or the technological evolution in general, additional works could explore further driving and retarding factors.

In terms of the *survival analysis*, especially *because* the corporate landscape seems to signal that nanotechnology – at least in Germany – seems to be rather "bottom-up"- than "top-down"-driven with large chemical companies playing a dominant role, further variables could for instance capture specific developments within the chemical industry and relate them to the development of nanotechnology. Also, size- or regional effects could be subject to further investigation.

Regarding the *social network analysis*, examining the networks of technological overlap could be particularly rewarding for firms: Due to the apparent relation between the *technological diversity* and *technological survival* of firms, companies could – for example – look out for an ideal company to cooperate with or to acquire in order to enlarge their patenting portfolio and thereby to enhance their (technological) survival chances in the area of nanotechnology. For scientific works it could then be specifically interesting to examine actual cooperations between firms. Inspecting patent applications in this respect may serve as one possible approach, but also, boards of directors of different companies could be investigated to determine whether the detected technological proximity (or distance) of companies can be confirmed and explained by personal affiliations. In a further step, the effects of a firm's network of relations on innovation could for example be assessed following AHUJA (2000). Networks of personal relations as in CANTNER AND GRAF (2006) could be constructed to gain further insight into interfirm relations.

To detect whether the discovered patterns are rather unique to the domestic nanotechnology environment or pertain to other technologies and countries as well, the presented survival and social network analysis could furthermore be carried out for diverse technologies and/ or diverse countries. Having detected potential differences, in a second step, possibly, strengths and weaknesses in the domestic nanotechnology environment could be revealed. Finally, these differences could be related to technology- and/ or country-specific framework conditions, such as legislation, subsidies and/ or other backgrounds in order to identify possible starting points for improvements.

In conclusion, this work provides two empirical concepts for explaining the evolution of domestic nanotechnology companies. It presents tools which are amendable and scalable so that they can be adjusted and refined further to even better capture the (future) characteristics and development of the corporate domestic nanotechnology landscape. Eventually, such a tool can be used to support and foster the *sustainable* development of domestic nanotechnology companies and thereby contribute to the country's international attractiveness and competitiveness.

5 APPENDIX

5.1 Appendix to section 1

5.1.1 Summary of risks for mankind and for the environment

In section 1.1, chances and risks associated with nanotechnology are mentioned. While remarks on chances are rather comprehensive, remarks on risks are rather confined to the statement that they are identified for mankind and for the environment. This section serves to describe the risks associated with nanotechnology in more detail.

Concerning the risks for mankind, three types of exposition are distinguished: the absorption of nanoparticles over inhalation, over oral and over dermal exposition (see KÜHLING AND HORN 2007: 11FF). With regards to the *inhalation* of nanoparticles, a study performed on rats showed that nanoparticles are able to enter the animal's brain after being inhaled over the animal's nose. At the same time the study points out that possible negative effects are not sufficiently examined (see Oberdörster 2005 cit. after UMWELTBUNDESAMT 2006: 13). Regarding the *oral* exposition only few studies seem to be present so that a toxicological evaluation is difficult; in general toxicity seems to depend on the local and systemic distribution (see KÜHLING AND HORN 2007: 13). Also, concerning the *dermal* exposition, at this point, a final assessment seems impossible: Surveys on sunscreens for example find out that (given intact skin) nanoparticles do not seem to enter the human body if they are larger than 20nm (see Wiench 2006 cit. after KÜHLING AND HORN 2007: 15). However, according to Baroli, injured skin may let nanoparticles pass into subjacent layers of skin (see ZENTRUM DER GESUNDHEIT 2009). In summary, at this point the exploration of risks for mankind is fragmentary, so a final evaluation is still pending.

A similar picture is revealed concerning the environmental risks: In terms of the fauna, there are several studies concerning the exposition of nanoparticles to fishes. In one of the studies fishes are exposed to nC_{60} – the so-called fullerenes. Amongst other, the findings are that the fishes suffer a significant lipid peroxidation after 48 hours of exposure to 0.5 ppm uncoated nC_{60} (see OBERDÖRSTER 2004: 1058). However, a later study points out that the *preparation* of nC_{60} (e.g. by tetrahydrofuran (THF)) gravely influences the toxicity of nC_{60} (see OBERDÖRSTER ET AL. 2005: 1112). Other environmental risks concern the ground, waters and air. In this respect, there seem only few works available. One work for example depicts a reduced root growth when exposing crop plants to aluminum nanoparticles (see KÜHLING AND HORN 2007: 15). Altogether, as in case of the risks for mankind, environmental risks are largely unexplored.

5.1.2 BMBF affirmative actions

In connection with the chances and risks associated with nanotechnology, section 1.1 outlines the efforts which are undertaken to further push nanotechnology forward. Next to initiatives on European level, BMBF expenditures on nanotechnology are mentioned to have quadrupled in the time between 1998 and 2004. Figure 40 (below) provides some more details on the BMBF expenditures including the main areas which are fostered.

Figure 40: Expenditures for nanotechnology within the scope of diverse BMBF key issues.

Nanotechnologie-förderung des BMBF (in Mio. EUR)	**Schwerpunkte**	**1998**	**2002**	**2003**	**2004**	**2005**
Nanomaterialien	Nanoanalytik, Nanobiotechnologie, Nanostrukturmaterialien, Nanochemie, CCN, Nanonachwuchswettbewerb, Nanochance		19,2	20,3	32,7	38,1
Produktionstechnologien	Ultradünne Schichten, ultrapräzise Oberflächen		0,2	0,8	2,2	2,2
Optische Technologien	Nanooptik, Ultrapräzisionsbearbeitung, Mikroskopie, photonische Kristalle, Molekularelektronik, Diodenlaser, OLED		18,5	25,2	26	26
Mikrosystemtechnik	Systemintegration		7	7	9,4	10,2
Kommunikations-technologien	Quantenstruktursysteme, photonische Kristalle		4,3	4	3,6	3,4
Nanoelektronik	EUVL, Lithografie, Maskentechnologie, eBiochips, Magnetoelektronik, SiGe-Elektronik,		19,9	25	44,7	46,2
Nanobiotechnologie	Manipulationstechniken, funktionalisierte Nanopartikel, Biochips,		4,6	5,4	5	3,1
Innovations- und Technikanalysen	ITA-Studien		0,2	0,5	0,2	
Summe (in Mio. EUR)		**27.6**	**73,9**	**88,2**	**123,8**	**129,2**

Source: ZUKÜNFTIGE TECHNOLOGIEN CONSULTING DER VDI TECHNOLOGIEZENTRUM GMBH (2004: 32).

A large share of expenditures is thereby spent on nano electronics. Also nano materials are strongly fostered. Altogether, the sum of expenditures has increased significantly.

5.2 Appendix to section 2

5.2.1 Keyword list

Section 2.3.2 concentrates on the sources used for the determination of nanotechnology companies. Altogether, five sources are used. One of these sources is the "Förderkatalog"[125]. Following the database operator, company data from this catalogue is retrieved by searching for the term "nano" in the database containing BMBF/ BMWi-fostered projects. The refined resulting list (shown in section 2.3.2, paragraph (1)) includes projects financed by the BMBF or BMWi from the budget of nanotechnology. Nanotechnology projects which are not financed from this budget are not included in the list. Table 29 (below) depicts, how such projects could be retrieved from the "Förderkatalog": The presented keyword list was configured by the VDI to deduce nanotechnology projects from the "Förderkatalog" and used by HEINZE 2006 (see HEINZE 2006: 281-283).

Table 29: Keyword list.

%abform%mikro%	%euv%
%afm%	%funktion%supramol%
%asphär%	%halbleiter%
%bauelement%elektronik%	%ionen%ultra%
%biophotonik%	%katalys%mikro%
%detektion%mikro%	%katalys%oxidat%
%diode%laser%	%katalysator%mikro%reaktion%
%elektron%strukturierung%	%keramik%sensor%
%elektronik%bauelement%	%keramisch%sensor%
%euv%lithographie%	%kompos%nano%
%germanium%silizium%	%mikro%abform%
%ion%lithografie%	%mikro%katalys%
%ion%litographie%	%mikro%reaktion%katalys%
%kosmetik%	%mram%
%kristall%photon%	%nano%komp%
%laser%diode%	%nano%kontakt%
%lithografie%ion%	%nanokomposit%
%lithographie%ion%	%oxidat%katalys%
%magnetoelektronik%	%oxidation%katalys%
%mikro%detektion%	%präzis%schicht%
%mikro%sonde%	%replikation%

[125] BUNDESMINISTERIUM FÜR BILDUNG UND FORSCHUNG/ BUNDESMINISTERIUM FÜR WIRTSCHAFT UND TECHNOLOGIE: Förderkatalog, http://oas2.ip.kp.dlr.de/foekat/foekat/foekat, 16 January 2007.

%nano%	%schicht%präzis%
%nanoanalytik%	%schicht%ultradünn%
%nanometer%	%sensor%keramik%
%photon%kristall%	%sensor%keramisch%
%schicht%solar%	%strukturier%technik%
%silizium%germanium%	%strukturierung%elektron%
%solar%schicht%	%supramolekular%
%sonde%mikro%	%technik%strukturier%
	%ultradünn%schicht%

Source: HEINZE (2006: 282).

However, since – due to the evolving and strongly diversified area of nanotechnology – the keyword list is not up to date, from a present point of view, the additional usage of this keyword list seems obsolete.

5.2.2 Search strategy

For the completion of dataset 1 used for the survival analysis (see section 2.3.3) as well as for dataset 2 (see section 3.3.2) used for the social network analysis, patent data is needed. In this book, patent data is retrieved from PATSTAT (VERSION 10/ 2007) using a specific search strategy. Table 30 (below) delivers the complete search strategy which was developed by Fraunhofer, ISI, in: NOYONS ET AL. (2003: 100-101). The search strategy contains a list of search strings (each indicated by the letter “S”) which is being searched for, in case of this book in PATSTAT (VERSION 10/ 2007). The strings consist of a combination of keywords, which are often followed by a specific punctuation and/ or letters. The punctuations (“#”, “!” and “?”) and letters (“W” and “2A”) are assigned a particular role. Their role is explained by the legend included at the end of the table.

Table 30: Search strategy for patent applications in nanotechnology.

S (((NANOMETER# OR NANOMETRE# OR NM OR SUBMICRO?) AND (CHIP# OR ELECTRON? OR ENGINEERING OR DIAMETER OR SIZE# OR LAYER# OR SCALE OR ORDER OR RANGE OR DIMENSIONAL))/TI NOT (WAVELENGTH# OR ROUGHNESS OR ABSORB?)/TI)
S (((NANOMETER# OR NANOMETRE# OR NM OR SUBMICRO?)(A)(CHIP# OR ELECTRON? OR ENGINEERING OR DIAMETER OR SIZE# OR LAYER# OR SMALL? OR SCALE OR ORDER OR RANGE OR DIMENSIONAL)) NOT (WAVELENGTH# OR ROUGHNESS OR ABSORB?))
S (((NANOMETER# OR NANOMETRE# OR NM OR SUBMICRO?)(2W)(CHIP# OR

ELECTRON? OR
ENGINEERING OR DIAMETER OR SIZE# OR LAYER# OR SMALL? OR SCALE OR ORDER OR
RANGE OR DIMENSIONAL)) NOT (WAVELENGTH# OR ROUGHNESS OR ABSORB?))
S (NANOPARTICL? OR NANO(W)PARTICL?) NOT (ABSORB? OR INK OR POLISH?)
S (NANOANALY? OR NANOBAR? OR NANOBOT# OR NANOCAGE# OR NANOCHANNEL? OR
NANOCERAMIC OR NANOCHANNEL# OR NANOCHIP# OR NANOCIRCUITRY OR
NANOCLUSTER# OR NANOCOATING# OR NANOCOLL? OR NANOCOMPUT? OR
NANOCOMPOS? OR NANOCONDUCT? OR NANOCRY OR NANOCRYSTAL? OR NANODEVICE#
OR NANODES)
S (NANODIMENSIONAL OR NANODISPERS? OR NANODOMAIN# OR NANODROP? OR
NANOENGIN? OR NANOELECTR? OR NANOFABRIC? OR NANOFEATURE# OR NANOARRAY?
OR NANOBIO? OR NANOREACT? OR NANOCATAL? OR NANOPHOTO? OR NANOHOL? OR
NANOPIT# OR NANOPILLAR#)
S (NANOGAP# OR NANOGEL OR NANOGLASS? OR NANOGRAIN? OR NANOGRANULAR OR
NANOGRID? OR NANOIMPRINT? OR NANOINDENTATION OR NANOINSTRUCTIONS OR
NANOILLUMINATION)
S (NANOLAYER? OR NANOLITHO? OR NANOMACHIN? OR NANOMANIPULATOR# OR
NANOMAGNET? OR NANOMATERIAL?)
S (NANOMECHANICAL OR NANOMEMBRANE OR NANOMETRIC? OR NANOMICR? OR
NANOMOTOR# OR NANOPEPTID? OR NANOPHASE# OR NANOPHOTOLITHOGRAPHY OR
NANOPIPEL? OR NANOPLOTTER# OR NANOPOWDER# OR NANOSENSOR# OR NANOSCALE?
OR NANOARCHITECTURE OR NANOPATTERN OR NANOCAVITIY)
S (NANOPOR? OR NANOPRINTING OR NANOPROBES OR NANOPROCESS? OR
NANOPROGRAM? OR NANORIBBONS OR NANOROD# OR NANOROPE# OR NANOSCIEN? OR
NANOSCOP? OR NANOSCRATCHING OR NANOSEMICONDUCTOR# OR NANOSENS? OR
NANOSEQUENCER OR NANOSILIC? OR NANOSILVER OR NANOSIZ?)
S (NANOSPHER? OR NANOSPREADING OR NANOSTATS OR NANOSTEP? OR NANOSTRUCT?
OR NANOSUBSTRATE OR NANOSUSPENSION OR NANOSWITCH? OR NANOSYST? OR

NANOTECHNOLOG? OR NANOTEXTUR? OR NANOTIPS OR NANOTRIBOLOGY OR NANOTROPES OR NANOTUB? OR NANOWIRE? OR NANOWHISK?)
S (NANOTOPOGRAPHY OR NANOCHEMISTRY OR NANOREGOGNITION OR NANO-DOT OR
NANOPUMP# OR NANOCAPS?)
S SCANNING PROBE MICROSCOP? OR SCANNING TUNNEL? MICROSCOP? OR SCANNING
FORCE MICROSCOP? OR ATOMIC FORCE MICROSCOP? OR NEAR FIELD MICRO-SCOP?
S FUNCTIONALLY COATED SURFACE# AND NANO?
S (BIOCHIP OR BIOSENSOR) AND (A61# OR G01N OR C12Q)/IC
S DNA(W)CMOS
101
S (BACTERIORHODOPSIN OR BIOPOLYMER# OR BIOMOLECULE#)AND (G11# OR G02# OR
G03# OR G06#)/IC
S BIOMOLECULAR TEMPLAT? OR VIRUS(2A)ENCAPSULATION OR MODIFIED VI-RUS
S NANO? AND IMPLANT?
S (PATTERN? OR ORGANIZED) AND (BIOCOMPATABILITY OR BLOODCOMPATABIL-ITY OR
BLOOD COMPATABILITY OR CELL SEEDING OR CELLSEEDING OR CELL THERAPY OR
TISSUE REPAIR OR EXTRACELLULAR MATRIX OR TISSUE ENGINEERING OR BIO-SENSOR#
OR IMMUNOSENSOR# OR BIOCHIP OR CELL ADHESION)
S MICRO?(2A)NANO?
S NANO(W)(ARCHITECT? OR CERAMIC OR CLUSTER# OR COATING# OR COMPO-SIT## OR
CRYSTAL?)
S NANO(W)(DEVICE# OR DISPERSE# OR DIMENSIONAL OR DISPERSION# OR DROP# OR
DROPLET OR ENGINEERING OR ENGINEERED OR ELECTRODES OR ELECTRON-IC#)
S NANO(W)(FABRICATED OR FABRICATION OR FILLER# OR GEL OR GRAIN? OR IMPRINT
OR IMPRINTED OR LAYER#)
S NANO(W)(MACHINE# OR MANIPULATOR# OR MATERIAL# OR MECHANICAL OR MEMBRANE OR METRIC?)
S NANO(W)(PHASE# OR POWDER# OR PORE# OR PORO? OR PRINTING OR ROD# OR SCALAR)
S NANO(W)(SIZE? OR SPHER# OR STRUCTURE# OR STRUCTURING OR SUSPEN-SION OR
SYSTEM# OR TECHNOLOG?)

S NANO(W)(TEXTUR? OR TIPS OR TROPES OR TUB? OR WIRE? OR WHISK?)
S ATOMIC(W)LAYER# OR MOLECULAR TEMPLATES OR SUPRAMOLECULAR CHE-MISTRY
OR MOLECULAR MANIPULATION
S QUANTUM DEVICE# OR QUANTUM DOT# OR LANGMUIR BLODGETT OR QUAN-TUM WIRE?
S SINGLE ELECTRON? TUNNELING OR MOLECUL? ENGINEER? OR MOLECUL? MANUFACTUR?
S MOLECUL? SELF ASSEMBL? OR ULTRAVIOLET LITHOGRAPHY OR PDMS STAMP OR SOFT
LITHOGRAPHY
S FULLEREN? OR MOLECULAR MOTOR OR MOLECULAR BEACON OR NANO ELECTROSPRAY OR ION CHANNELS OR MOLECULE CHANNELS
S LAB(3W)CHIP
S (NANOFILT? OR NANOFIB? OR NANOFLUID?) AND (C0## OR A61# OR B0##)/IC
S (ELECTRON BEAM WRITING) AND (H01L OR H01J)/IC
S MONOLAYER AND (G03G OR H01J)/IC
S THIOL AND H01L/IC
S (B82B OR A61K009-51 OR G01N013-10 OR G12B021)/IC
S L1-L39

Note:
truncation up to one character (0 or 1)
! truncation of exactly one character (1)
? unlimited truncation (0 or any number)
W directly adjacent terms in order specified
2A adjacent terms in any order, separated by up to 2 words

Source: Fraunhofer, ISI, in: NOYONS ET AL. (2003: 100-101).

5.2.3 Sources listing application fields

To describe dataset 1, amongst other, in section 2.3.4, application fields of nanotechnology companies are under investigation (see paragraph (1)). Since there is no common source listing application fields of companies in a consistent manner, NACE-codes (see section 5.2.4) are employed in section 2. To provide an impression of the heterogeneity of sources listing application fields, the following two subsections yield some examples of sources listing application fields of companies. Section 5.2.3.1 depicts two sources which list application fields of a selection of nanotechnology companies. Section 5.2.3.2 lists six sources which name application fields of companies which are referred to as biotechnology companies in the source but in part appear to be nanotechnology companies according to the sources used in this book.

5.2.3.1 General

- VDI-TECHNOLOGIEZENTRUM: Nanomap, http://www.nano-map.de/, 7 August 2007.
- HESSEN-NANOTECHNOLOGIE: KOMPETENZMATRIX UNTERNEHMEN 1, http://www.hessen-nanotech.de/mm/02_Kompetenzmatrix.pdf, 6 November 2008.

5.2.3.2 Biotechnology

- TECHNOLOGIEPARK HEIDELBERG (2004): Companies in the Field of Biotechnology BioRegion Rhein-Neckar-Dreieck, http://www.citebiotech.com/Docs/fra/Doc%20partenaires/Heidelberg%202004-09-29.pdf, 10 August 2009.
- GESELLSCHAFT FÜR BIOANALYTIK MÜNSTER E. V. (2004): Münster biotech-region companies and competence, http://www.bioanalytik-muenster.de/conpresso/_data/Biotechregion_3-1.pdf, 10 August 2009.
- BIOM AG MUNICH BIOTECH DEVELOPMENT (2005): Die BioTech-Region München – Erfolg schafft neue Zuversicht, http://www.bio-m.org/_resources/dynamic/hauptbereich/bio_m_publikationen/report_2005_deutsch.pdf, 10 August 2009.
- BIOTOP BERLIN-BRANDENBURG (2008): Biotech-Report 2007·2008, Ausgabe 34 Mai 2008, http://www.biotop.de/data/files/biotopics/Branchenreport2008-D.pdf, 10 August 2009.
- LANDESENTWICKLUNGSGESELLSCHAFT (LEG) THÜRINGEN MBH ABTEILUNG AKQUISITION UND INTERNATIONALE KONTAKTE (2004): Biotechnologie in Thüringen, http://www.leg-thueringen.de/fileadmin/www/pdfs/DE/publikationen/factsheet_bio.pdf, 10 August 2009.
- BIORIVER – LIFE SCIENCE IM RHEINLAND E.V. (2007): BioRiver Report 2007, http://www.bioriver.de/fileadmin/downloads/BioRiver_Report_Deu.pdf, 10 August 2009.

5.2.4 NACE

Following section 2.3.4 and section 5.2.3, in this book, NACE-codes are used to retrieve information about the branches the companies are active in. Such NACE-codes consist of a key (a two-digit number) and its respective label (the textual translation of the key). Table 31 (below) shows the complete list of NACE-codes or NACE-keys provided by HOPPENSTEDT (2009: 1-10)[126] which are referred to this book.

[126] According to HOPPENSTEDT (2009), the applied nomenclature is based on the source “Klassifikation der Wirtschaftszweige 2008 (WZ 2008) mit Erläuterungen” published by the Statistisches Bundesamt, Wiesbaden 2009.

Table 31: NACE-keys.

Key	Label
01	Landwirtschaft, Jagd und damit verbundene Tätigkeiten
02	Forstwirtschaft und Holzeinschlag
03	Fischerei und Aquakultur
05	Kohlenbergbau
06	Gewinnung von Erdöl und Erdgas
07	Erzbergbau
08	Gewinnung von Steinen und Erden, sonstiger Bergbau
09	Erbringung von Dienstleistungen für den Bergbau und für die Gewinnung von Steinen und Erden
10	Herstellung von Nahrungs und Futtermitteln
11	Getränkeherstellung
12	Tabakverarbeitung
13	Herstellung von Textilien
14	Herstellung von Bekleidung
15	Herstellung von Leder, Lederwaren und Schuhen
16	Herstellung von Holz-, Flecht-, Korb- und Korkwaren (ohne Möbel)
17	Herstellung von Papier, Pappe und Waren daraus
18	Herstellung von Druckerzeugnissen; Vervielfältigung von bespielten Ton-, Bild- und Datenträgern
19	Kokerei und Mineralölverarbeitung
20	Herstellung von chemischen Erzeugnissen
21	Herstellung von pharmazeutischen Erzeugnissen
22	Herstellung von Gummi und Kunststoffwaren
23	Herstellung von Glas und Glaswaren, Keramik, Verarbeitung von Steinen und Erden
24	Metallerzeugung und -bearbeitung
25	Herstellung von Metallerzeugnissen
26	Herstellung von Datenverarbeitungsgeräten, elektronischen und optischen Erzeugnissen
27	Herstellung von elektrischen Ausrüstungen
28	Maschinenbau
29	Herstellung von Kraftwagen und Kraftwagenteilen
30	Sonstiger Fahrzeugbau
31	Herstellung von Möbeln
32	Herstellung von sonstigen Waren
33	Reparatur und Installation von Maschinen und Ausrüstungen
35	Energieversorgung
36	Wasserversorgung
37	Abwasserentsorgung

Key	Label
38	Sammlung, Behandlung und Beseitigung von Abfällen; Rückgewinnung
39	Beseitigung von Umweltverschmutzungen und sonstige Entsorgung
41	Hochbau
42	Tiefbau
43	Vorbereitende Baustellenarbeiten, Bauinstallation und sonstiges Ausbaugewerbe
45	Handel mit Kraftfahrzeugen; Instandhaltung und Reparatur von Kraftfahrzeugen
46	Großhandel (ohne Handel mit Kraftfahrzeugen)
47	Einzelhandel (ohne Handel mit Kraftfahrzeugen)
49	Landverkehr und Transport in Rohrfernleitungen
50	Schifffahrt
51	Luftfahrt
52	Lagerei sowie Erbringung von sonstigen Dienstleistungen für den Verkehr
53	Post-, Kurier- und Expressdienste
55	Beherbergung
56	Gastronomie
58	Verlagswesen
59	Herstellung, Verleih und Vertrieb von Filmen und Fernsehprogrammen; Kinos; Tonstudios und Verlegen von Musik
60	Rundfunkveranstalter
61	Telekommunikation
62	Erbringung von Dienstleistungen der Informationstechnologie
63	Informationsdienstleistungen
64	Erbringung von Finanzdienstleistungen
65	Versicherungen, Rückversicherungen und Pensionskassen (ohne Sozialversicherung)
66	Mit Finanz- und Versicherungsdienstleistungen verbundene Tätigkeiten
68	Grundstücks und Wohnungswesen
69	Rechts- und Steuerberatung, Wirtschaftsprüfung
70	Verwaltung und Führung von Unternehmen und Betrieben; Unternehmensberatung
71	Architektur- und Ingenieurbüros; technische, physikalische und chemische Untersuchung
72	Forschung und Entwicklung
73	Werbung und Marktforschung
74	Sonstige freiberufliche, wissenschaftliche und technische Tätigkeiten
75	Veterinärwesen
77	Vermietung von beweglichen Sachen
78	Vermittlung und Überlassung von Arbeitskräften
79	Reisebüros, Reiseveranstalter und Erbringung sonstiger Reservierungsdienstleistungen

Key	Label
80	Wach- und Sicherheitsdienste sowie Detekteien
81	Gebäudebetreuung; Garten- und Landschaftsbau
82	Erbringung von wirtschaftlichen Dienstleistungen für Unternehmen und Privatpersonen a. n. g.
84	Öffentliche Verwaltung, Verteidigung; Sozialversicherung
85	Erziehung und Unterricht
86	Gesundheitswesen
87	Heime (ohne Erholungs- und Ferienheime)
88	Sozialwesen (ohne Heime)
90	Kreative, künstlerische und unterhaltende Tätigkeiten
91	Bibliotheken, Archive, Museen, botanische und zoologische Gärten
92	Spiel-, Wett- und Lotteriewesen
93	Erbringung von Dienstleistungen des Sports, der Unterhaltung und der Erholung
94	Interessenvertretungen sowie kirchliche und sonstige religiöse Vereinigungen (ohne Sozialwesen und Sport)
95	Reparatur von Datenverarbeitungsgeräten und Gebrauchsgütern
96	Erbringung von sonstigen überwiegend persönlichen Dienstleistungen
97	Private Haushalte mit Hauspersonal
98	Herstellung von Waren und Erbringung von Dienstleistungen durch private Haushalte für den Eigenbedarf ohne ausgeprägten Schwerpunkt
99	Exterritoriale Organisationen und Körperschaften

Source: HOPPENSTEDT (2009: 1-10).[127]

5.2.5 Kaplan-Meier estimates: computation of mean survival times

In connection with the Kaplan-Meier estimates (see section 2.5.1), the mean survival time of companies is computed in general and per covariate, for example for companies with pre-entry experience and for those without. To compute the mean survival time, two cases need to be distinguished because else, the resulting mean survival time might be biased: If the last observation is not censored (case 1), the mean survival time can be computed employing the restricted mean. The "[...] rmean (*restricted mean*) computes the mean survival time restricted to the longest follow up time." (STATACORP LP 2007: 114). It equals the area under the Kaplan-Meier survival function (see STATACORP LP 2007: 116). If the last observed analysis time (in this book: time since entry) is censored (case 2, which is always the case in this book), the restricted mean underestimates the true mean (see CLEVES ET AL. 2008: 120). To circumvent making this mistake, STATA yields the option of using the *extended mean*. "emean (extended mean) computes the mean survival time by exponentially extending the survival curve to

[127] According to HOPPENSTEDT (2009), the applied nomenclature is based on the source "Klassifikation der Wirtschaftszweige 2008 (WZ 2008) mit Erläuterungen" published by the Statistisches Bundesamt, Wiesbaden 2009.

zero [...]" (STATACORP LP 2007: 114). Due to the artificial approximation, the emean does not equal the true mean either – this should be considered in the interpretation. However, the extended mean appears closer to reality than the restricted mean, which is why the extended mean is used to compute the mean survival times in section 2.

5.2.6 Testing methods of the assumption of proportionality

In connection with the choice of the most suitable regression model (see section 2.4.2.1), two criteria are mentioned: Criterion one concerns the behavior of the hazard over time, criterion two concerns the supposition of the hazard to follow a specific distribution. While in terms of the latter, the main text contains rich information on how to determine the potential distribution of the hazard, in terms of the behavior of the hazard over time, remarks are kept rather short. It is mentioned, that the assumption of proportionality has to be investigated for its validity. In this context, three methods are named: Two of these methods ("stphplot" and "stcoxkm") are graphical evaluation methods, the third ("estat phtest") is a statistical evaluation method. To illustrate what these methods comprise, all three methods are briefly described in the following.

The stphplot-command "[...] plots –ln{–ln(survival)} curves for each category of a nominal or ordinal covariate versus ln(analysis time)[[128]]. [...] The proportional-hazards assumption is not violated when the curves are parallel." (STATACORP LP 2007: 158).

As another graphical evaluation tool, the stcoxkm on the other hand does the following (see STATACORP LP 2007: 158):

> "`stcoxkm` plots Kaplan-Meier observed survival curves and compares them with the Cox predicted curves for the same variable. The closer the observed values are to the predicted, the less likely it is that the proportional-hazards assumption has been violated. […]"

As the interpretation of graphical illustrations may depend on the reader, a statistical analysis to test whether the condition of proportionality is fulfilled can be achieved by the estat phtest. estat phtest "[…] tests proportional-hazards assumption on the basis of Schoenfeld residuals" (STATACORP LP 2007: 166). These have to be computed first before the actual test can take place. If multiple failures occur at the same time, the efron method needs to be applied. STATACORP LP (2007: 131) explains the efron method as follows:

> „[…] The Efron method […] for handling tied values assumes that the first risk pool is $e_1+e_2+e_3+e_4+e_5$ and the second risk pool is either $e_2+e_3+e_4+e_5$ or $e_1+e_3+e_4+e_5$. From this, Efron noted that the e_1 and e_2 terms were in the second risk pool with probability ½ and so used for the second risk pool $.5(e_1+e_2)+e_3+e_4+e_5$. Efron's approximation is a more accurate approximation

[128] In this book, the analysis time is referred to as the "time since entry". It is measured in years.

of the exact marginal likelihood than Breslow's but takes longer to calculate. […]"

In comparison to the graphical evaluation methods, estat phtest allows for the confirmation or the rejection of the assumption of proportionality by means of statistical significance. Since therefore, the estat phtest is not bound to a subjective assessment of the assumption of proportionality, in terms of the model choice (see section 2.5.2.1) the estat phtest is applied (see section 5.2.7.1).

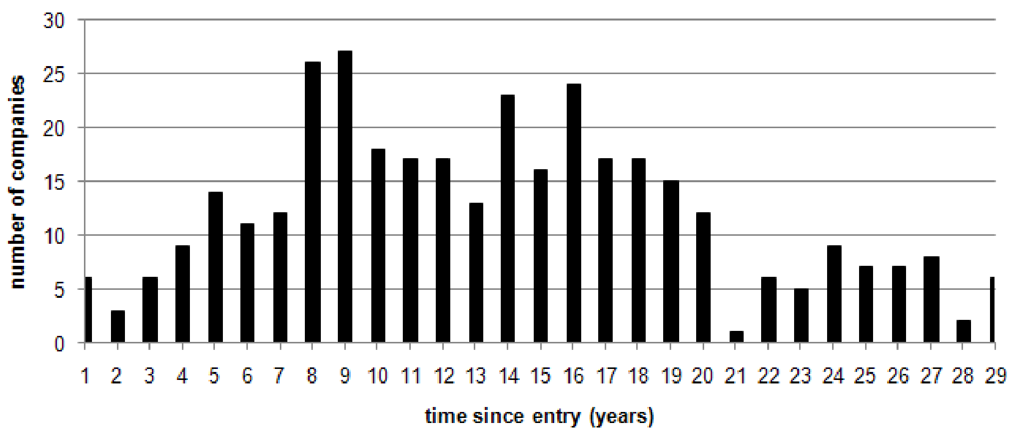
5.2.7 Regression analysis: model choice

Section 2.5.2.1 is concerned with the determination of the model best to apply in terms of dataset 1. Against this background, two steps are undertaken. They are described in the following two subsections (see section 5.2.7.1 and section 5.2.7.2).

5.2.7.1 Assumption of proportionality: results of the estat phtest

In order to choose the correct model as described in section 2.5.2.1, the assumption of proportionality has to be tested for its validity. This is achieved by the estat phtest (see section 5.2.6 for methodological remarks on the estat phtest). It is performed on dataset 1. To find the correct specification of the estat phtest, the test requires a preexamination of ties (multiple failures occurring at the same time) in the database. As the examination shows, multiple failures do occur in dataset 1 (see Figure 41 below).

Figure 41: Occurence of ties in dataset 1.



Source: Own compilations based on dataset 1.

For example, there are 27 study exits after 9 years, 18 study exits after 10 years on the market etc. As mentioned in section 2.5.2.2, to account for these ties, the efron method is applied which is introduced in section 5.2.6. The computation of the Schoenfeld residuals yields the results below (see Table 32 below).

Table 32: Cox regression (efron method for ties), Schoenfeld residuals.[129]

Log likelihood	-126.395
n	354
Prob>chi2	0.065
pre-entry experience	0.840 [0.727]
post-entry experience	
(cohort 1)	*(1.000)*
cohort 2	*3.500 [0.039]*
cohort 3	*5.723 [0.009]*
technological know-how	0.738 [0.653]

Source: Own compilations based on dataset 1.

The results of the estat phtest are depicted by Table 33 (below).

Table 33: Results of the estat phtest.

	rho	chi2	df	Prob>chi2
pre-entry experience	0.109	0.30	1	0.581
post-entry experience				
cohort 2	*-0.200*	*0.85*	*1*	*0.358*
cohort 3	*-0.418*	*2.84*	*1*	*0.092*
technological know-how	-0.410	4.31	1	0.038
global test		7.23	4	0.1240

Source: Own compilations based on dataset 1.

Following the estat phtest, there is no evidence that the specification violates the proportional hazards assumption: According to the global test, Prob>chi2 is not significant. Hence, findings point to the application of a PH-model to the data.

5.2.7.2 Behavior of the hazard

To determine whether the hazard follows a particular distribution, as described in section 2.5.2.1 two approaches following BRADBURN ET AL. (2006B: 608) are employed: one approach is based on the AIC of the models, the other approach is based on hazard estimates. The subsequent two subsections provide supple-

[129] Note that P-values are reported in parentheses [] besides the hazard rate.

mentary information on each and thereby serve to deliver more details than given in section 2.5.2.1.

(1) AIC for PH- and AFT-models

In case of the data underlying the examination (dataset 1), Table 34 (below) displays the AIC for the three PH-models and Table 35 (below) displays the AIC for the four computable AFT-models. It is important to note that all models include the covariates in their original form (“pre-entry experience”, “cohort 2”, “cohort 3” and “technological know-how”), so neither mathematical transformations nor stratification have/ has taken place. Since the covariate technological know-how displays a slight nonproportionality, this is especially to consider in terms of the PH-models.

Table 34: AIC for PH-models.

	Exponential	**Weibull**	**Gompertz**
log likelihood	-99.648	-96.409	-97.398
c	1	2	2
k	4	4	4
AIC	209.297	204.818	206.796

Source: Own compilations based on dataset 1 and CLEVES ET AL. (2008: 273).

Table 35: AIC for AFT-models.

	Exponential	**Weibull**	**Lognormal**	**Log-logistic**
log likelihood	-99.648	-96.409	-95.893	-96.315
c	1	2	2	2
k	4	4	4	4
AIC	209.297	204.818	203.787	204.629

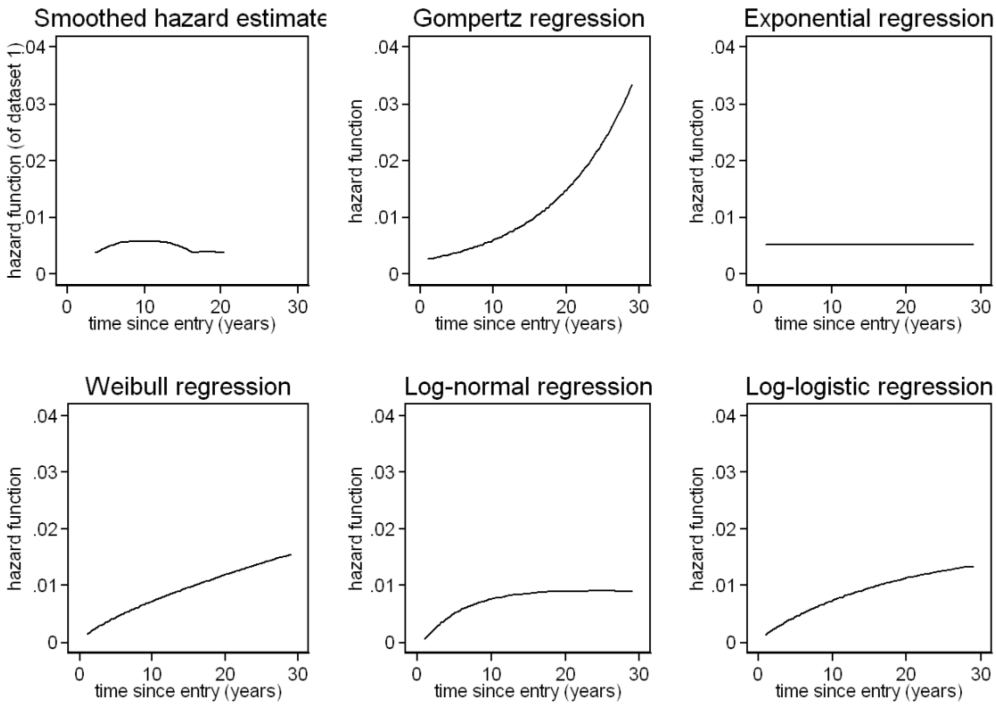
Source: Own compilations based on dataset 1 and CLEVES ET AL. (2008: 273).

According to Table 34, the Weibull distribution is the distribution best underlying the hazard in the sample. In case of the AFT-models depicted in Table 35, the Lognormal distribution would be another option. The Gamma-distribution did not deliver any results in terms of the underlying data.

(2) Hazard rates of the fitted fully parametric models

Next to the computation of the AIC of diverse models, following BRADBURN ET AL. (2006B: 608) the second approach to assess whether the hazard in the sample underlies a particular distribution is to plot the hazard rates of the fitted fully parametric approaches. Figure 42 (below) shows the results.

Figure 42: Hazard rates of the fitted fully parametric models.



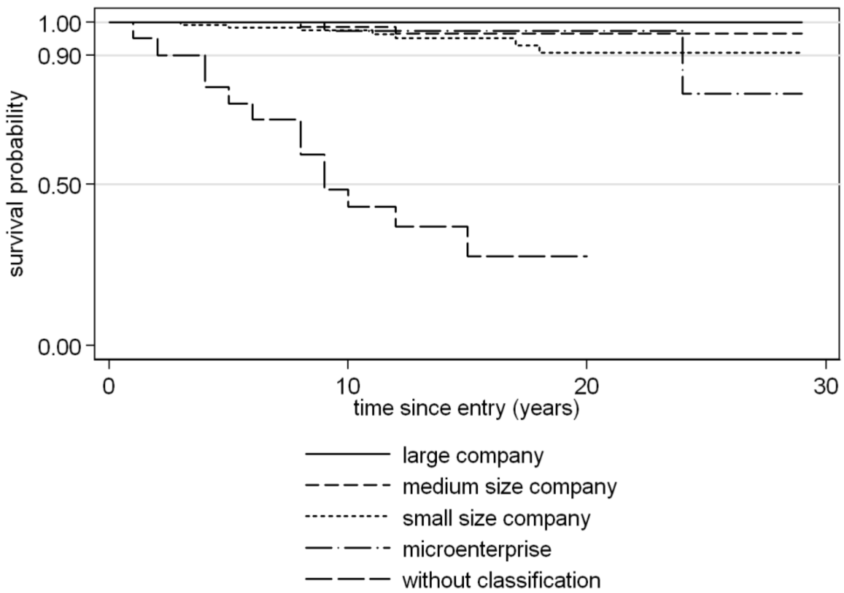
Source: Own compilations based on dataset 1.

As can be seen, the actual hazard (the “Smoothed hazard estimate”) differs from all other computed hazard rates. The Exponential model seems the most suitable distribution. The Weibull and the Lognormal model, both recommended by the AIC, clearly show deviations from the actual hazard. This is especially visible in later years, so deriving predictions on this basis seems rather critical. Altogether, this is regarded as a sign for the fully parametric models not to be adequate, so they are not considered in this book in more detail.

5.2.8 Kaplan-Meier curves for the company size

In connection with the summary and conclusion of section 2 provided in section 2.6, potential improvements of the survival analysis are addressed. One aspect made to the subject of discussion is the inclusion of size effects of companies. Figure 43 (below) shows the Kaplan-Meier curves for the diverse company sizes.

Figure 43: Kaplan-Meier curves for the company size.



Source: Own compilations based on dataset 1.

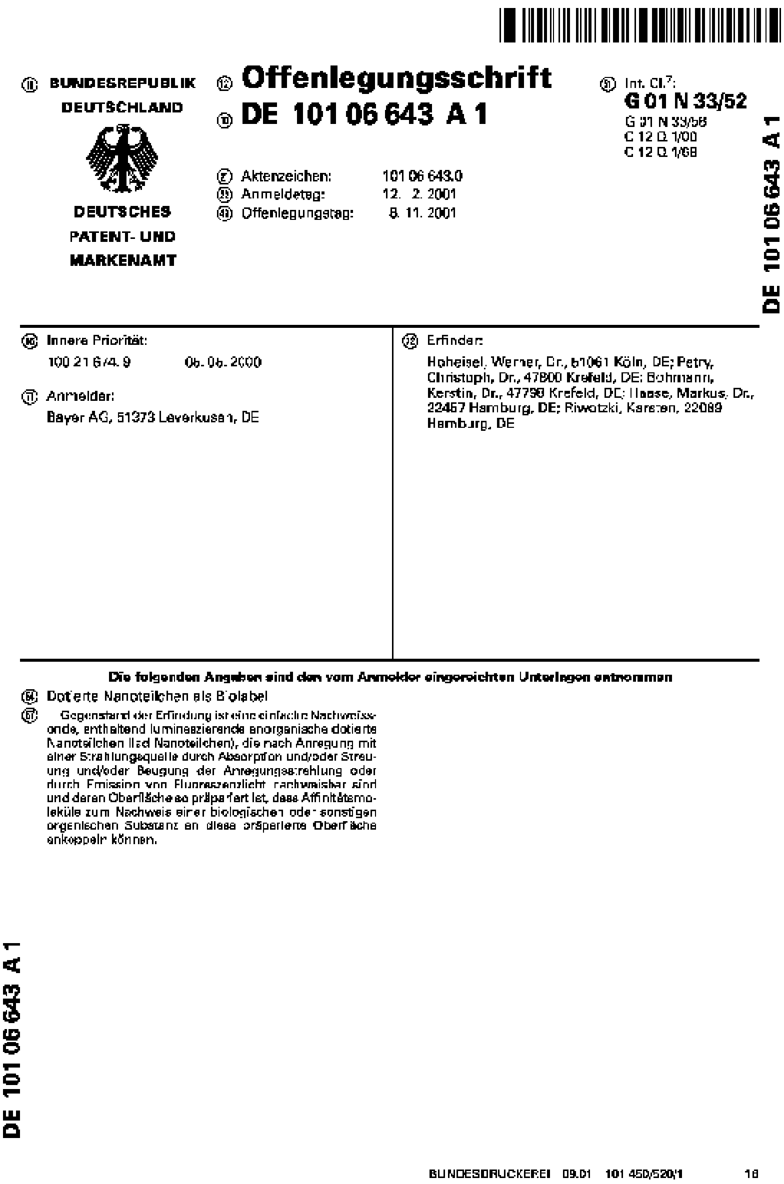
Originally, size effects were thought to be included in the survival analysis. However, as can be deduced from Figure 43, measured by the extreme decrease of the step function, the share of exiting companies is high amongst the companies without classification. Excluding the respective companies from the analysis would have been equal to reducing the number of exits in the sample to a critically low extent which would have impaired the explanatory power of the models. This guided the decision of disregarding size effects in the survival analysis.

5.3 Appendix to section 3

5.3.1 Exemplary patent application

In section 3, technological fields listed on nanotechnology patent applications are linked to the technological persistence of companies in the area of nanotechnology. To perform such an analysis, suitable data needs to be at hand. This data is available on patent (applications). To give an idea, which kind of data is included on patent (applications), Figure 44 (below) provides an example.

Figure 44: Exemplary patent application at the DPMA.



(19) BUNDESREPUBLIK DEUTSCHLAND

DEUTSCHES PATENT- UND MARKENAMT

(12) **Offenlegungsschrift**

(10) **DE 101 06 643 A 1**

(51) Int. Cl.[7]:
G 01 N 33/52
G 01 N 33/58
C 12 Q 1/00
C 12 Q 1/68

(21) Aktenzeichen: 101 06 643.0
(22) Anmeldetag: 12. 2. 2001
(43) Offenlegungstag: 8. 11. 2001

DE 101 06 643 A 1

(66) Innere Priorität:
100 21 674. 9 06. 06. 2000

(71) Anmelder:
Bayer AG, 51373 Leverkusen, DE

(72) Erfinder:
Hoheisel, Werner, Dr., 51061 Köln, DE; Petry, Christoph, Dr., 47800 Krefeld, DE; Bohmann, Kerstin, Dr., 47798 Krefeld, DE; Haase, Markus, Dr., 22457 Hamburg, DE; Riwotzki, Karsten, 22089 Hamburg, DE

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen

(54) Dotierte Nanoteilchen als Biolabel

(57) Gegenstand der Erfindung ist eine einfache Nachweissonde, enthaltend lumineszierende anorganische dotierte Nanoteilchen (lad Nanoteilchen), die nach Anregung mit einer Strahlungsquelle durch Absorption und/oder Streuung und/oder Beugung der Anregungsstrahlung oder durch Emission von Fluoreszenzlicht nachweisbar sind und deren Oberfläche so präpariert ist, dass Affinitätsmoleküle zum Nachweis einer biologischen oder sonstigen organischen Substanz an diese präparierte Oberfläche ankoppeln können.

DE 101 06 643 A 1

BUNDESDRUCKEREI 09.01 101 450/520/1 16

Source: DEPATISNET[130], *DE000010106643A1.*

[130] DEPATISNET: http://depatisnet.dpma.de, 5 January 2011.

From the patent application it can for example be deduced that the applicant is Bayer AG and that the patent application has a priority year of 2000. Furthermore, the patent application has been assigned to several IPC classes, amongst other C12Q 1/00. Transferred into technological fields by the help of WIPO IPC-Technology Concordance Table (see SCHMOCH 2008: 9-10 or section 5.3.4), this would relate to technological field 15 which is Biotechnology. The technological fields chosen by companies are then related to the technological survival of companies, i.e. their technological presence or absence in the subsequent cohort.

5.3.2 JAVA programs

Section 3.3.3 is concerned with the preparation of the patent data (dataset 2) in a manner so that a social network analysis can be performed on it. Starting from EXCEL spreadsheets, the data needs to be transformed into a two-mode sociomatrix and then into an adjacency matrix. The following two Java programs serve to perform these two transformations.

5.3.2.1 Program 1 "build matrix"

The following program constructs a two-mode sociomatrix on the basis of dataset 2 (only patent data is thereby considered in which a domestic company occurs). The rows of the matrix contain the company names, the columns of the matrix contain the technological fields. Whenever a company touches upon a technological field over its patent applications, the respective position in the matrix contains the entry "1", else the respective position in the matrix contains the entry "0". Potentially empty fields can be manually removed from the analysis by marking the respective checkbox in the occuring window. It is also possible to choose between priority and publication year and to adjust for the time spans to analyze. Having computed the two-mode sociomatrix, in a second step, the programm computes the transposed two-mode sociomatrix. Both matrices are stored in a separate CSV-file in preparation of the further analysis. The program was written by Sebastian Schirp, a former student assistant at the Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW), on behalf of the author of this book. The general procedure of constructing a two-mode sociomatrix is taken from CANTNER AND GRAF (2006: 466-477).

```
import java.awt.*;
import java.awt.event.*;
import javax.swing.*;
import java.io.*;
import java.util.*;
import java.util.regex.*;
import java.nio.charset.*;


public class Build_Matrix extends JFrame
{
      private class MeinWindowListener extends WindowAdapter
```

```
	{
		public void windowClosing(WindowEvent arg0)
		{
		System.exit(0);
	}
	}
	private JComponent contentPane = (JComponent)
getContentPane();
	private JButton button;
	private JRadioButton[] choice=new JRadioButton[2];
	private ButtonGroup bg=new ButtonGroup();
	private JLabel a=new JLabel("Input-Dateipfad");
	private JLabel b=new JLabel("Ausgabedateipfad");
	private JTextArea e1=new JTextArea(".csv");
	private JTextArea e2=new JTextArea("");
	private JLabel c= new JLabel("Anfangsjahr");
	private JLabel d= new JLabel("Endjahr");
	private JCheckBox box=new JCheckBox();
	private JLabel e= new JLabel("#nv mit einbeziehen");
	private JTextArea e3=new JTextArea("");
	private JTextArea e4=new JTextArea("");
	private JPanel panel=new JPanel();
	private JPanel panel2=new JPanel();
	private JPanel buttons=new JPanel();


	private String eingang=new String();
	private String ausgang=new String();
	private int anfjahr=0;
	private int endjahr=0;
	private int auswahl=3;
	private boolean nv=false;
	public ArrayList id;
	public ArrayList ipc;


	Build_Matrix() //Konstruktor - Fenster
	{
	super("Build Matrix");
	setSize(800, 300);
	initLayout();
	setLocationRelativeTo(null);
	addWindowListener(new MeinWindowListener());
		setVisible(true);
	setResizable(true);
	}



	public static void main(String[] args)
	{
		Build_Matrix frame = new Build_Matrix();
	frame.setLocation(150, 90);
	frame.setResizable(false);
	frame.setVisible(true);
	}

	void initLayout() //Layout
	{
```

```
            contentPane.setLayout(new BorderLayout());
            choice[0]=new JRadioButton("PrioritätsJahr", true);
            choice[1]=new JRadioButton("PublikationsJahr");
            bg.add(choice[0]);
            bg.add(choice[1]);
      button=new JButton("START");
      buttons.setLayout(new FlowLayout(FlowLayout.CENTER));
      buttons.setVisible(true);
      buttons.add(button);
      contentPane.add("South", buttons);

      panel.setLayout(new FlowLayout(FlowLayout.LEFT));
      panel.setVisible(true);
      panel.add(a);
      e1.setPreferredSize(new Dimension(680, 20));
      panel.add(e1);
      panel.add(b);
      e2.setPreferredSize(new Dimension(670, 20));
      panel.add(e2);
      panel.add(c);
      e3.setPreferredSize(new Dimension(50, 20));
      panel.add(e3);
      panel.add(d);
      e4.setPreferredSize(new Dimension(50, 20));
      panel.add(e4);

      panel.add(e);
      panel.add(box);
      panel.add(choice[0]);
      panel.add(choice[1]);

      contentPane.add("Center", panel);


            button.addActionListener(new ActionListener()
        {
                  public void actionPerformed(ActionEvent e)
//Knopffunktion
                  {

                        eingang=(e1.getText().trim());
                  ausgang=(e2.getText().trim());
                  anfjahr=Integer.parseInt(e3.getText().trim());
                  endjahr=Integer.parseInt(e4.getText().trim());
                  if (choice[1].isSelected())
                        auswahl=4;
                  else
                        auswahl=3;
                  nv=box.isSelected();
                  start();
                  }

            });
      }


      public void start ()
      {
```

```
            id=new ArrayList();
            ipc=new ArrayList();
            initList();
            buildTableOne();

      }


      public void initList()
      {
            try
            {
                  BufferedReader br=new BufferedReader(new
FileReader(eingang));
                  br.readLine();
                  while (br.ready())
                  {
                        String[]
dummy=br.readLine().replaceAll("\"","").split(";");
                        if
(anfjahr<=Integer.parseInt(dummy[auswahl])&&Integer.parseInt(dummy
[auswahl])<=endjahr&&Integer.parseInt(dummy[33])==1&&dummy[15].tri
m().equals("DE"))
                        {
                              if
(nv==true||!dummy[9].equalsIgnoreCase(""))
                              {
                                    if (!id.contains(dummy[12]))
                                          id.add(dummy[12]);
                                    if (!ipc.contains(dummy[9]))
                                          ipc.add(dummy[9]);
                              }
                        }
                        else
                              System.out.println(dummy[auswahl]+"
      "+dummy[15]+"     "+dummy[33]);

                  }
                  br.close();
            }
            catch(Exception e){System.out.println(e.toString());}
      }

      public void buildTableOne()
      {
            try
            {
                  BufferedReader br=new BufferedReader(new
FileReader(eingang));
                  BufferedWriter bw=new BufferedWriter(new
FileWriter(new File(ausgang+".csv")));
                  PrintWriter pw=new PrintWriter(bw);
                  br.readLine();
                  byte[][] matrix=new byte
[id.size()][ipc.size()];
                  while (br.ready())
                  {
                        String[]
dummy=br.readLine().replaceAll("\"","").split(";");
```

```
                        if
(anfjahr<=Integer.parseInt(dummy[auswahl])&&Integer.parseInt(dummy
[auswahl])<=endjahr&&Integer.parseInt(dummy[33])==1&&dummy[15].tri
m().equals("DE")&&(nv==true||!dummy[9].equalsIgnoreCase("")))
                        {

      matrix[id.indexOf(dummy[12])][ipc.indexOf(dummy[9])]=1;
                        }
                  }
                  for (int i=0;i<ipc.size();i++)
                        pw.print(";"+ipc.get(i).toString());
                  pw.println();
                  for (int i=0;i<id.size();i++)
                  {
                        pw.print(id.get(i).toString());
                        for (int j=0; j<ipc.size();j++)
                              pw.print(";"+matrix[i][j]);
                        pw.println();
                  }
                  br.close();
                  bw.close();
                  pw.close();
                  bw=new BufferedWriter(new FileWriter(new
File(ausgang+"_transp.csv")));
                  pw=new PrintWriter(bw);
                  for (int i=0;i<id.size();i++)
                        pw.print(";"+id.get(i).toString());
                  pw.println();
                  for (int i=0;i<ipc.size();i++)
                  {
                        pw.print(ipc.get(i).toString());
                        for (int j=0; j<id.size();j++)
                              pw.print(";"+matrix[j][i]);
                        pw.println();
                  }
                  bw.close();
                  pw.close();
            }
            catch(Exception e){System.out.println(e.toString());}
      }


}
```

5.3.2.2 Program 2: "multiply matrix"

The program below constructs an adjacency matrix on the basis of the previously computed two-mode sociomatrix and the transposed two-mode sociomatrix (see section 5.3.2.1). The entries in the matrix reflect the strength of the technological connection between the diverse companies. Three files are created: one NET-file including the matrix diagonal, one NET-file without diagonal and one CSV-file. The latter two are included for informative reasons, they are not needed for further analysis. In turn, the NET-file (including the matrix diagonal) is intended for further analysis, it is in Pajek-processible format. The program was written by Sebastian Schirp, a former student assistant at the Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW), on behalf of the author of this book. The general procedure of constructing an adjacency matrix is taken from CANTNER AND GRAF (2006: 466-477).

```
import java.awt.*;
import java.awt.event.*;
import javax.swing.*;
import java.io.*;
import java.util.*;
import java.util.regex.*;
import java.nio.charset.*;


public class Multiply_Matrix_axbt extends JFrame
{
      private class MeinWindowListener extends WindowAdapter
      {
            public void windowClosing(WindowEvent arg0)
            {
            System.exit(0);
      }
      }
      private JComponent contentPane = (JComponent)
getContentPane();
      private JButton button;
      private JLabel a=new JLabel("Input-Dateipfad A");
      private JLabel c=new JLabel("Output-Dateipfad");
      private JTextArea e1=new JTextArea(".csv");
      private JTextArea e3=new JTextArea("");
      private JPanel panel=new JPanel();
      private JPanel buttons=new JPanel();
      private JLabel ausgabeLab=new JLabel("hier stehen evtl.
Fehlermeldungen");

      private String eingang1=new String();
      private String eingang2=new String();
      private String ausgang=new String();

      public Integer [][] MatrA;
      public Integer [][] MatrB;
      public Integer [][] MatrC;
      public String[] BeschrZ;
      public String[] BeschrS;

      Multiply_Matrix_axbt() //Konstruktor - Fenster
      {
      super("Multiply Matrix");
      setSize(800, 300);
      initLayout();
      setLocationRelativeTo(null);
      addWindowListener(new MeinWindowListener());
            setVisible(true);
      setResizable(true);
      }
```

```
public static void main(String[] args)
{
Multiply_Matrix_axbt frame = new Multiply_Matrix_axbt();
frame.setLocation(150, 90);
frame.setResizable(false);
frame.setVisible(true);
}

void initLayout() //Layout
{
      contentPane.setLayout(new BorderLayout());

button=new JButton("START");
buttons.setLayout(new FlowLayout(FlowLayout.CENTER));
buttons.setVisible(true);
buttons.add(button);
contentPane.add("South", buttons);

panel.setLayout(new FlowLayout(FlowLayout.LEFT));
panel.setVisible(true);
panel.add(a);
e1.setPreferredSize(new Dimension(680, 20));
panel.add(e1);

panel.add(c);
e3.setPreferredSize(new Dimension(680, 20));
panel.add(e3);

contentPane.add("Center", panel);
contentPane.add("North", ausgabeLab);

      button.addActionListener(new ActionListener()
  {
            public void actionPerformed(ActionEvent e)
//Knopffunktion
            {

                  eingang1=(e1.getText());

                  ausgang=(e3.getText());
                  start();
            }
      });
}


public void start ()
{
      getMatrices();
      multiply();
      printOut();
}

public void getMatrices()
{
      try
      {
```

```
                  BufferedReader brA=new BufferedReader(new
FileReader(eingang1));
                  int columnsA= (brA.readLine().split(";").length-
1);
                  int rowsA=0;
                  while (brA.ready())
                  {
                        brA.readLine();
                        rowsA++;
                  }
                  BufferedReader brB=new BufferedReader(new
FileReader(eingang1));
                  int columnsB= (brB.readLine().split(";").length-
1);
                  int rowsB=0;
                  while (brB.ready())
                  {
                        brB.readLine();
                        rowsB++;
                  }

                  brB=new BufferedReader(new
FileReader(eingang1));
                  brA=new BufferedReader(new
FileReader(eingang1));  //reader1 neu erstellen um wieder am
anfang zu sein
                  brA.readLine();
                  brB.readLine();
                  MatrA=new Integer [rowsA][columnsA];
                  MatrB=new Integer [columnsB][rowsB];
                  MatrC=new Integer [rowsA][rowsB];
                  int counter=0;
                  BeschrZ=new String[rowsA];
                  BeschrS=new String[rowsB];
                  while (brA.ready())
                  {
                        String [] dummy=brA.readLine().split(";");
                        BeschrZ[counter]=dummy[0];
                        for (int i=1;i<dummy.length;i++)
                        {
                              MatrA [counter][i-
1]=Integer.parseInt(dummy[i]);
                        }
                        counter++;
                  }
                  counter=0;
                  while (brB.ready())
                  {
                        String [] dummy=brB.readLine().split(";");
                        BeschrS[counter]=dummy[0];
                        for (int i=1;i<dummy.length;i++)
                        {
                              MatrB [i-
1][counter]=Integer.parseInt(dummy[i]);
                        }
                        counter++;
                  }
                  for(int i=0;i<MatrC.length;i++)
                        for (int j=0;j<MatrC[i].length;j++)
```

```
                        MatrC[i][j]=0;
        }
        catch (Exception e){System.out.println(e.toString());}
    }

    public void  multiply()
    {
        for(int i=0;i<MatrC.length;i++)
            for (int j=0;j<MatrC[i].length;j++)
                for (int k=0;k<MatrA [0].length;k++)

    MatrC[i][j]+=(MatrA[i][k]*MatrB[k][j]);
    }

    public void printOut()
    {
        try
        {
            BufferedWriter bw=new BufferedWriter(new
FileWriter(new File(ausgang+"_pajek.net")));
            PrintWriter pw=new PrintWriter(bw);
            pw.println("*Vertices "+BeschrS.length);
            for (int j=1;j<=BeschrS.length;j++)
                pw.println(j+" \""+BeschrS[j-1]+"\"");
            pw.println("*Matrix");
            for (int i=0;i<MatrC.length;i++)
            {
                for (int j=0;j<MatrC[i].length;j++)
                    pw.print(MatrC[i][j]+"\t");
                pw.println();
            }
            bw.close();
            pw.close();

            bw=new BufferedWriter(new FileWriter(new
File(ausgang+"_pajek_no_diag.net")));
            pw=new PrintWriter(bw);
            pw.println("*Vertices "+BeschrS.length);
            for (int j=1;j<=BeschrS.length;j++)
                pw.println(j+" \""+BeschrS[j-1]+"\"");
            pw.println();
            pw.println("*Matrix");
            for (int i=0;i<MatrC.length;i++)
            {
                for (int j=0;j<MatrC[i].length;j++)
                    if (i!=j)
                        pw.print(MatrC[i][j]+"\t");
                    else
                        pw.print("0"+"\t");
                pw.println();
            }
            bw.close();
            pw.close();

            bw=new BufferedWriter(new FileWriter(new
File(ausgang+"view.csv")));
            pw=new PrintWriter(bw);
            for (int j=0;j<BeschrS.length;j++)
                pw.print(";"+BeschrS[j]);
```

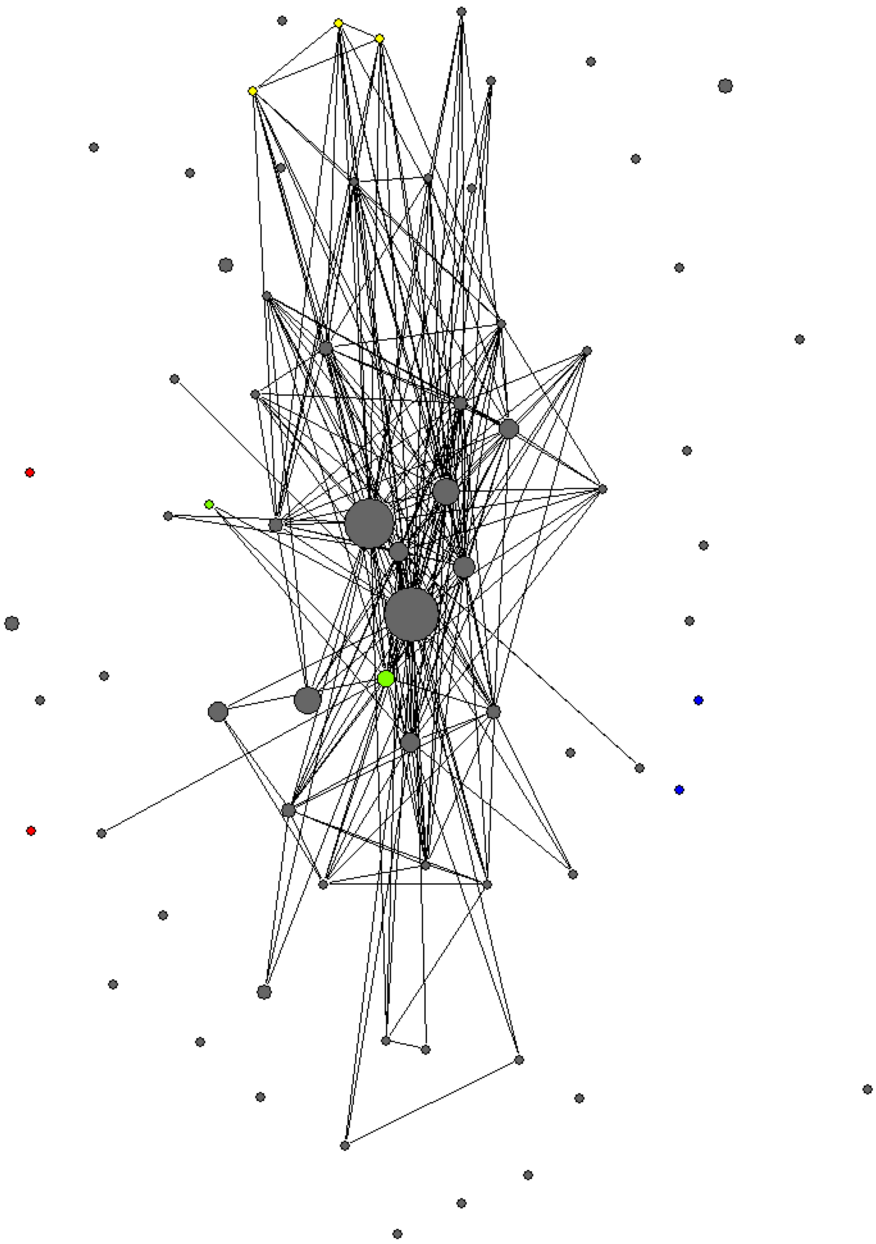
```
                    pw.println();
                    for (int i=0;i<MatrC.length;i++)
                    {
                          pw.print(BeschrS[i]);
                          for (int j=0;j<MatrC[i].length;j++)
                                if (j!=i)
                                      pw.print(";"+MatrC[i][j]);
                                else
                                      pw.print(";"+"-");
                          pw.println();
                    }
                    bw.close();
                    pw.close();

              }
              catch(Exception e){System.out.println(e.toString());}
        }
}
```

5.3.3 Cooperation in the networks of technological overlap

In section 3.3.4, in connection with the remarks on applicants given in paragraph (4) and in section 3.6, the occurrence and respectively, increase of cooperations between companies is referred to. This section serves to deliver a little more insight on such cooperations.

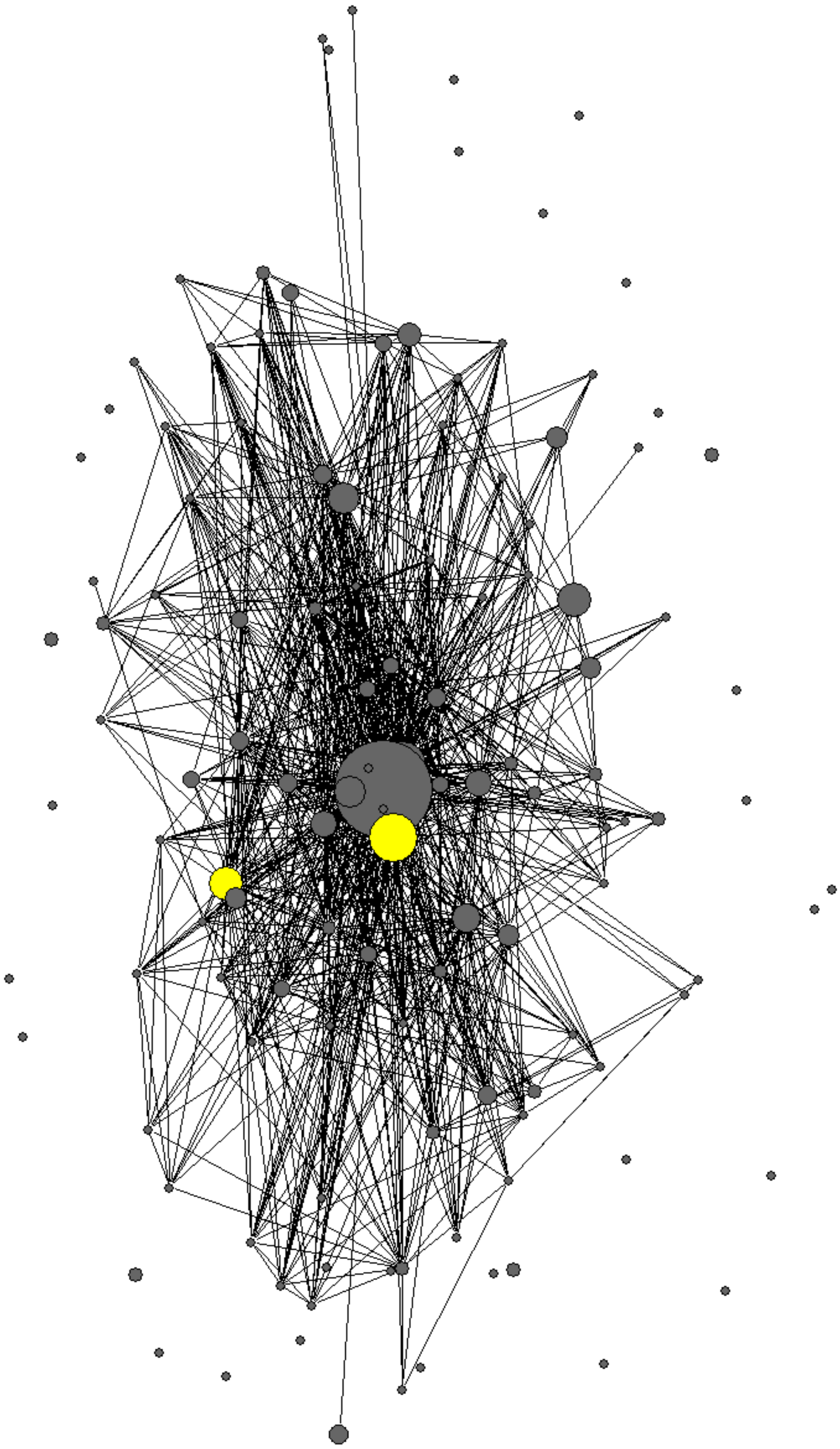
Figure 45, Figure 46 and Figure 47 (below) show the few *actual* cooperations (measured using patent applications) in the networks of technological overlap. Since the companies in cohort 1 and cohort 2 do not cooperate, only the networks for cohort 3, 4 and 5 are presented. In the figures, the grey vertices depict non-cooperating companies and the colored vertices depict the diverse cooperations. In general, each color demarks a different cooperation. It is to remark that in cohort 4 the depicted companies have two patent applications in common. In cohort 5, there are companies which cooperate with more than one company in various patents. In this case, the same coloring of the vertices is applied, so equal colors do not necessarily imply that all companies cooperate on one patent. As before, to ease readability, only lines with a line value of at least two are considered.

Figure 45: Cooperation in the network of technological overlap, cohort 3.



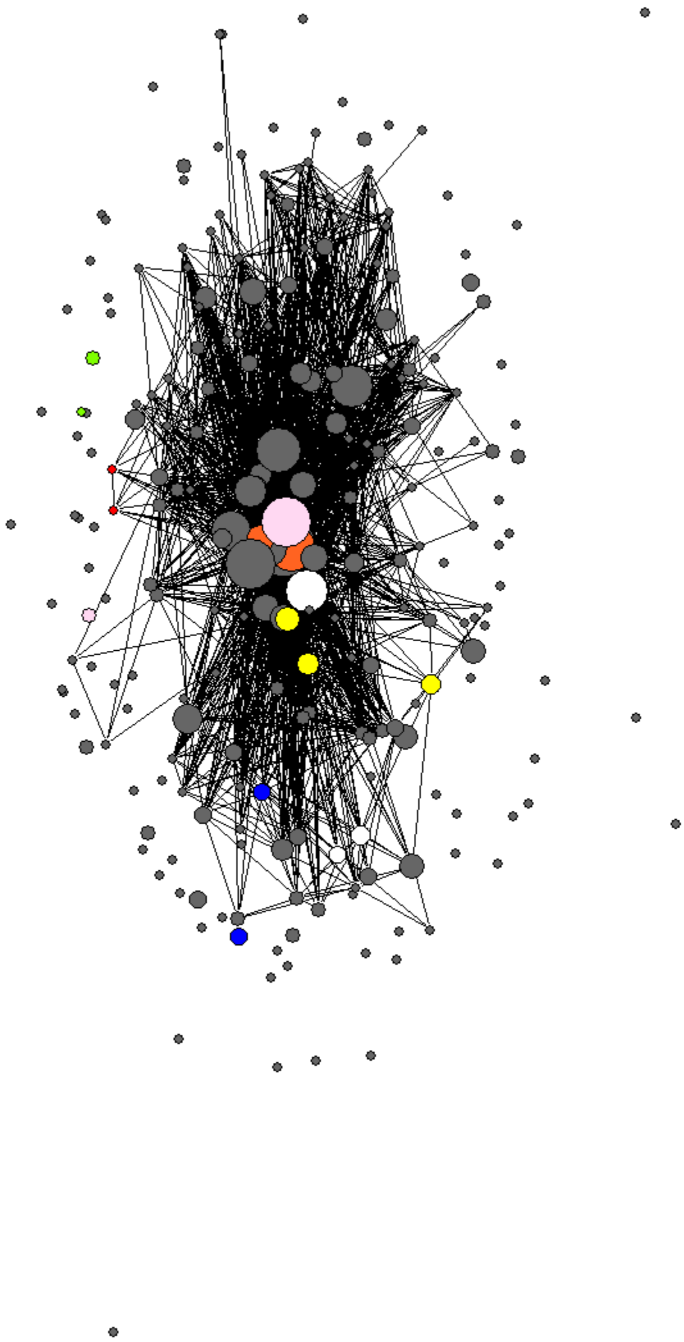
Source: Own computations based on dataset 2.

Figure 46: Cooperation in the network of technological overlap, cohort 4.



Source: Own computations based on dataset 2.

Figure 47: Cooperation in the network of technological overlap, cohort 5.



Source: *Own computations based on dataset 2.*

In cohort 3, two of four cooperations are “outside” of the core, which – for reasons of simplicity – is considered the cluster of highly connected companies in the center of the network. Interestingly, between none of the cooperating companies, a specific relation, e.g. in the sense of parent company – subsidiary, could be determined. Instead, it is striking that the cooperating companies have a rather low number of patent applications.

In cohort 4, the two cooperating companies are inside the core. One of the companies thereby originated from the other, but was sold to a group of investors prior to their two common patent applications. It is furthermore remarkable that compared to other companies, in general, both companies have a higher number of patent applications each.

Cohort 5 differs from the remaining cohorts in various ways: On the one hand, in cohort 5, an increasing number of cooperations is present. On the other hand, most of the cooperations occur in the core of the networks. When taking a detailed look at the cooperations taking place, it becomes obvious that, in this cohort, relations between companies are manifold. It occurs that one company is founded by the other company. One firm has entered a cooperation treaty with another company with the aim of developing a specific common product. The same firm has entered a partnership with a second company. The remaining companies either seem to have no particular relationship, or one company is part of the other. Altogether, it is striking that cooperating companies usually apply for more patents.

5.3.4 WIPO IPC-Technology Concordance Table

Section 3.3.4, paragraph (7) is concerned with the derivation of technological fields from the IPC classes referred to by patent applications. To do so, the classification scheme as recommended by SCHMOCH 2008 is used. The classification – called WIPO IPC-Technology Concordance Table – distinguishes amongst 35 fields of technology (see SCHMOCH 2008: 9-10). Figure 48 depicts these 35 fields of technology.

Figure 48: WIPO IPC-Technology Concordance Table.

	Field of Technology	International Patent Classification (IPC) Symbols
I: Electrical engineering		
1	Electrical machinery, apparatus, energy	F21#, H01B, H01C, H01F, H01G, H01H, H01J, H01K, H01M, H01R, H01T, H02#, H05B, H05C, H05F, H99Z
2	Audio-visual technology	G09F, G09G, G11B, H04N-003, H04N-005, H04N-009, H04N-013, H04N-015, H04N-017, H04R, H04S, H05K
3	Telecommunications	G08C, H01P, H01Q, H04B, H04H, H04J, H04K, H04M, H04N-001, H04N-007, H04N-011, H04Q
4	Digital communication	H04L
5	Basic communication processes	H03#
6	Computer technology	(G06# not G06Q), G11C, G10L
7	IT methods for management	G06Q
8	Semiconductors	H01L
II: Instruments		
9	Optics	G02#, G03B, G03C, G03D, G03F, G03G, G03H, H01S
10	Measurement	G01B, G01C, G01D, G01F, G01G, G01H, G01J, G01K, G01L, G01M, (G01N not G01N-033), G01P, G01R, G01S; G01V, G01W, G04#, G12B, G99Z
11	Analysis of biological materials	G01N-033
12	Control	G05B, G05D, G05F, G07#, G08B, G08G, G09B, G09C, G09D
13	Medical technology	A61B, A61C, A61D, A61F, A61G, A61H, A61J, A61L, A61M, A61N, H05G
III: Chemistry		
14	Organic fine chemistry	(C07B, C07C, C07D, C07F, C07H, C07J, C40B) not A61K, A61K-008, A61Q
15	Biotechnology	(C07G, C07K, C12M, C12N, C12P, C12Q, C12R, C12S) not A61K
16	Pharmaceuticals	A61K not A61K-008
17	Macromolecular chemistry, polymers	C08B, C08C, C08F, C08G, C08H, C08K, C08L
18	Food chemistry	A01H, A21D, A23B, A23C, A23D, A23F, A23G, A23J, A23K, A23L, C12C, C12F, C12G, C12H, C12J, C13D, C13F, C13J, C13K
19	Basic materials chemistry	A01N, A01P, C05#, C06#, C09B, C09C, C09F, C09G, C09H, C09K, C09D, C09J, C10B, C10C, C10F, C10G, C10H, C10J, C10K, C10L, C10M, C10N, C11B, C11C, C11D, C99Z
20	Materials, metallurgy	C01#, C03C, C04#, C21#, C22#, B22#
21	Surface technology, coating	B05C, B05D, B32#, C23#, C25#, C30#
22	Micro-structural and nano-technology	B81#, B82#
23	Chemical engineering	B01B, B01D-000#, B01D-01##, B01D-02##, B01D-03##, B01D-041, B01D-043, B01D-057, B01D-059, B01D-06##, B01D-07##, B01F, B01J, B01L, B02C, B03#, B04#, B05B, B06B, B07#, B08#, D06B, D06C, D06L, F25J, F26#, C14C, H05H
24	Environmental technology	A62D, B01D-045, B01D-046, B01D-047, B01D-049, B01D-050, B01D-051, B01D-052, B01D-053, B09#, B65F, C02#, F01N, F23G, F23J, G01T, E01F-008, A62C
IV: Mechanical engineering		
25	Handling	B25J, B65B, B65C, B65D, B65G, B65H, B66#, B67#
26	Machine tools	B21#, B23#, B24#, B26D, B26F, B27#, B30#, B25B, B25C, B25D, B25F, B25G, B25H, B26B
27	Engines, pumps, turbines	F01B, F01C, F01D, F01K, F01L, F01M, F01P, F02#, F03#, F04#, F23R, G21#, F99Z
28	Textile and paper machines	A41H, A43D, A46D, C14B, D01#, D02#, D03#, D04B, D04C, D04G, D04H, D05#, D06G, D06H, D06J, D06M, D06P, D06Q, D99Z, B31#, D21#, B41#
29	Other special machines	A01B, A01C, A01D, A01F, A01G, A01J, A01K, A01L, A01M, A21B, A21C, A22#, A23N, A23P, B02B, C12L, C13C, C13G, C13H, B28#, B29#, C03B, C08J, B99Z, F41#, F42#
30	Thermal processes and apparatus	F22#, F23B, F23C, F23D, F23H, F23K, F23L, F23M, F23N, F23Q, F24#, F25B, F25C, F27#, F28#
31	Mechanical elements	F15#, F16#, F17#, G05G
32	Transport	B60#, B61#, B62#, B63B, B63C, B63G, B63H, B63J, B64#
IV: Other fields		
33	Furniture, games	A47#, A63#
34	Other consumer goods	A24#, A41B, A41C, A41D, A41F, A41G, A42#, A43B, A43C, A44#, A45#, A46B, A62B, B42#, B43#, D04D, D07#, G10B, G10C, G10D, G10F, G10G, G10H, G10K, B44#, B68#, D06F, D06N, F25D, A99Z
35	Civil engineering	E02#, E01B, E01C, E01D, E01F-001, E01F-003, E01F-005, E01F-007, E01F-009, E01F-01#, E01H, E03#, E04#, E05#, E06#, E21#, E99Z

Source: WIPO IPC-TECHNOLOGY CONCORDANCE TABLE.

In the table, the column on the right depicts the IPC classes on subclass level, the column in the middle and the left column depict the technological fields which are associated with the respective IPC subclasses. Furthermore, the technological fields are summarized into five main groups.

5.3.5 Closeness- and betweenness-based measurement concepts

Section 3.4.2 is concerned with the identification and implication of (de-) centralized networks. In this respect, three measurement concepts are mentioned: the degree-, the closeness- and the betweenness-based measurement concept. Due to its suitability and usage in terms of this book, subsection 3.4.2.1 provides detailed remarks on the degree-based measurement concept. In consequence of their subordinate role they are assigned in this book, the other two measurement concepts are not described in detail. For reasons of competition and comparability and to yield an alternative view of centrality and centralization, the following two sections contain remarks on the closeness- and betweenness-based measurement concept.

5.3.5.1 Closeness-based measurement concept

As the degree-based measurement concept, also the closeness-based measurement concept consists of two actor-based measures (the closeness centrality and the normalized closeness centrality) and one network-based measure (closeness centralization). The measures are presented in consecutive order.

According to JANSEN (2006: 137) the *closeness centrality* of an actor n_i ($C_C(n_i)$) captures the nearness of the actor to all other actors in the network using distances[131]. Mathematically, this is expressed as

$$C_C(n_i) = \left[\sum_{j=1}^{n} d(n_i, n_j)\right]^{-1} \text{ and } i \neq j$$

where n is the total number of network members and $d(n_i,n_j)$ is the distance between actors n_i and n_j, or, in other words, the number of edges between n_i and n_j (see JANSEN 2006: 137). Further following JANSEN (2006: 137), $d(n_i,n_j)$ is only computable when the actors are connected, otherwise it is not determinable.

In connected networks, at maximum, an actor is directly connected to all other actors, so the distance to each other actor is 1; in sum this yields a distance of (n-1). This results into a maximum closeness centrality of 1/(n-1). If an actor is only distantly related to all other actors, in sum, this implies a higher distance and hence leads to a lower closeness centrality value. In the example network A (see p. 81, Figure 19) n_5 is connected to all other actors. It has a maximum closeness centrality of $C_C(n_5)=1/4$. The remaining actors have a closeness centrality of $C_C(n_1)=C_C(n_2)=C_C(n_3)=C_C(n_4)=1/7$. In network B, $C_C(n_1)=C_C(n_2)=1/10$, $C_C(n_3)=C_C(n_4)=1/7$ and $C_C(n_5)=1/6$. Altogether, $C_C(n_i)$ therefore ranges in the interval (0, 1/(n-1)]. The upper interval boundary shows that the measure closeness centrality is dependent on the network size. As in case of the measure degree centrality, therefore, when networks of different sizes are present, this makes comparisons between actors of different networks difficult.

To neutralize the effect of different network sizes, $C_C(n_i)$ is related to its maximal possible value of 1/(n-1) (see JANSEN 2006: 134, 137). This measure is referred to as the *normalized closeness centrality* of actor n_i ($C_{nC}(n_i)$). Mathematically, $C_{nC}(n_i)$ is computed according to the formula (see JANSEN 2006: 137)

---

[131] According to DE NOOY ET AL. (2005: 127), "[t]he *distance* from vertex u to vertex v is the length of the geodesic from u to v." Further following the author, "[a] geodesic is the shortest path between to vertices."

$$C_{nC}(n_i) = \frac{C_C(n_i)}{\frac{1}{n-1}} = \frac{n-1}{\sum_{j=1}^{n} d(n_i, n_j)}$$

It ranges in the interval (0, 1]. In example network A (see p. 79, Figure 19) normalized closeness centrality of $C_{nC}(n_5)$ is 1. The remaining actors have a normalized closeness centrality of $C_{nC}(n_1)=C_{nC}(n_2)=C_{nC}(n_3)=C_{nC}(n_4)=4/7$. In network B, $C_{nC}(n_1)=C_{nC}(n_2)=2/5$, $C_{nC}(n_3)=C_C(n_4)=4/7$ and $C_{nC}(n_5)=2/3$.

Based on the normalized closeness centrality of singular vertices, it is possible to compute a measure to evaluate the closeness centrality of an entire network, the so-called *closeness centralization* (C_C). DE NOOY ET AL. (2005: 127) define closeness centralization of a network as follows:

> "*Closeness centralization* is the variation in the closeness centrality of vertices divided by the maximum variation in closeness centrality scores possible in a network of the same size."

Mathematically, C_C is expressed as (see JANSEN 2006: 140)

$$C_C = \frac{\sum_{i=1}^{n}[C_{nC}(n^*) - C_{nC}(n_i)]}{\max \sum_{i=1}^{n}[C_{nC}(n^*) - C_{nC}(n_i)]} = \frac{\sum_i^n[C_{nC}(n^*) - C_{nC}(n_i)]}{\frac{(n^2-3n+2)}{(2n-3)}}$$

where $C_{nC}(n^*)$ is the normalized closeness centrality of the most central actor in the network and $C_{nC}(n_i)$ is the normalized closeness centrality of actor n_i. The numerator describes the variation in the normalized closeness centrality of the actors. The expression in the denominator again stands for the maximum possible network centralization (see JANSEN 2006: 140). This is achieved by a star network. In a star network, $C_{nC}(n_i)=(n-1)/(2n-3)$ pertains to all actors except for the central actor, who has a normalized closeness centrality of $C_{nC}(n^*)=1$. Concretely, the denominator $(n^2-3n+2)/(2n-3)$ is derived as follows:

$$\max \sum_{i=1}^{n}[C_{nC}(n^*) - C_{nC}(n_i)]$$

$$= (n-1) * \left[1 - \frac{n-1}{2n-3}\right] + 1 * [1-1] = \frac{n^2-3n+2}{2n-3}$$

$$\text{with } C_C(n^*) = \frac{1}{n-1}, C_C(n_i) = \frac{1}{2n-3}$$

Since the observed network is always related to the maximum possible network centralization, C_C ranges in the interval [0, 1]. A low value implies that – according to the closeness-based measure – a decentral network is present and a high value implies that a rather centralized network is present. In example network A (see p. 81, Figure 19) C_C=1, in network B, C_C=19/45.

By taking into account the "relatives" of actors, the measures seem rather sophisticated. However, they also come along with difficulties. As DE NOOY ET AL. (2005: 128) emphasize, if an undirected network is not connected (or analogously: a directed network is not strongly connected) implying that there are no paths between all vertices then distances between some vertices cannot be computed. Consequently, it is not possible to calculate closeness centrality for the vertices. In this case, the authors suggest considering only those vertices which are reachable to or from the vertex of interest and apply according weights. The authors furthermore portend that this solution does not work for computing the closeness centralization of the entire network. Unfortunately, this hinders comparison between networks. Furthermore, the measures seem rather unsuitable when being primarily concerned with the communication activity of a vertex or its ability to control communication within a network.

5.3.5.2 Betweenness-based measurement concept

The betweenness-based measurement concept also comprises two actor-based measures (the betweenness centrality and the normalized betweenness centrality) and one network-based measure (betweenness centralization). The measures are presented in consecutive order.

According to JANSEN (2006: 137) the *betweenness centrality* of an actor n_i ($C_B(n_i)$) captures the number of shortest connections (geodesics) between a pair of other actors which include actor n_i. With a slight modification[132] from JANSEN (2006: 137), mathematically, $C_B(n_i)$ is described as

$$\mathrm{C_B(n_i)} = \sum_{\mathrm{j}}^{\mathrm{n}} \sum_{\mathrm{k,k}>j}^{\mathrm{n}} \mathrm{b_{jk}(n_i)} \;\; \text{for } \mathrm{i \neq j \neq k}$$

$$\mathrm{b_{jk}(n_i)} = \frac{g_{jk}(n_i)}{g_{jk}}$$

where $g_{jk}(n_i)$ is the number of geodesics using n_i and g_{jk} is the number of geodesics between n_j and n_k. According to JANSEN (2006: 135) in other words, $b_{jk}(n_i)$ expresses the probability of a communication between two specific actors n_j and n_k to include actor n_i. n being the total number of network members, consequently, $C_B(n_i)$ is based on the probabilities that a communication of all actors n_j and all actors n_k includes actor n_i. Actor n_i should thereby not be part of the pair, which is expressed by $\mathrm{i \neq j \neq k}$. Else, still following JANSEN (2006: 135) the order of actors

[132] The index of "k" is written in a slightly modified manner.

is irrelevant – the consideration of [actor n_j, actor n_k] is sufficient which is why the instruction "j<k" is given. By doing so, double counting of the same pair of actors is avoided.

At maximum, an actor n_i is part of all shortest paths between all other pairs of actors. This is the case for the actor in the middle of a star network. In star networks, for the actor in the middle, $b_{jk}(n_i)$ is equal to one for each pair of actors n_j and n_k. $C_B(n_i)$ then is the number of unordered pairs in the network (which is n*(n-1)/2 according to JANSEN (2006: 136)) minus the number of pairs which include the actor in the middle (n-1). Mathematically, this leads a maximum possible betweenness centrality of $n*(n-1)/2-(n-1)=(n^2-3n+2)/2$. At minimum, an actor n_i is not part of any shortest paths between all other pairs of actors, so $b_{jk}(n_i)$ and, respectively, $C_B(n_i)$ are equal to zero. In example network A (see p. 79, Figure 19), n_5 lies on the shortest paths between (n_1, n_2), (n_1, n_3), (n_1, n_4), (n_2, n_3), (n_2, n_4), (n_3, n_4), so $C_B(n_5)=6$. For the remaining actors $C_B(n_1)=C_B(n_2)=C_B(n_3)=C_B(n_4)=0$. By contrast, in network B, $C_B(n_1)=C_B(n_2)=0$, $C_B(n_3)=C_B(n_4)=3$ and $C_B(n_5)=4$. Altogether, the measure betweenness centrality ranges in the interval of $[0, (n^2-3n+2)/2]$. The upper boundary shows that the measure is dependent on the network size. If networks of different sizes are present, this complicates comparisons between actors of different networks.

To neutralize the effect of different network sizes, $C_B(n_i)$ is related to its greatest possible value of $(n^2-3*n+2)/2)$ (see JANSEN 2006: 136-137). This measure is referred to as the *normalized betweenness centrality* of actor n_i ($C_{nB}(n_i)$) and is computed by following the mathematical expression (see JANSEN 2006: 137)

$$C_{nB}(n_i) = \frac{2C_B(n_i)}{n^2 - 3n + 2}$$

It ranges in the interval [0, 1]. In example network A (see p. 81, Figure 19), $C_B(n_5)=1$. For the remaining actors $C_B(n_1)=C_B(n_2)=C_B(n_3)=C_B(n_4)=0$. By contrast, in network B, $C_B(n_1)=C_B(n_2)=0$, $C_B(n_3)=C_B(n_4)=1/2$ and $C_B(n_5)=2/3$.

Using the normalized betweenness centrality, it is also possible to compute the betweenness centrality of an entire network, the so-called betweenness centralization (C_B). According to DE NOOY ET AL. (2005: 131) the betweenness centralization of a network is defined as follows:

> "*Betweenness centralization* is the variation in the betweenness centrality of vertices divided by the maximum variation in betweenness centrality scores possible in a network of the same size."

Mathematically, following JANSEN (2006: 141), C_B can be expressed as

$$C_B = \frac{\sum_{i=1}^{n}[C_{nB}(n^*) - C_{nB}(n_i)]}{\max \sum_{i=1}^{n}[C_{nB}(n^*) - C_{nB}(n_i)]} = \frac{\sum_{i=1}^{n}[C_{nB}(n^*) - C_{nB}(n_i)]}{(n-1)}$$

where $C_{nB}(n^*)$ is the normalized betweenness centrality of the most central actor in the network and $C_{nB}(n_i)$ is the normalized betweenness centrality of actor n_i. The numerator describes the variation in the normalized betweenness centrality of the actors. The expression in the denominator stands for the maximum possible network centralization which is achieved by a star network. In a star network, $C_{nB}(n_i)=0$ pertains to all actors except for the central actor, who has a normalized betweenness centrality of $C_{nB}(n^*)=1$. Concretely, the denominator (n-1) is derived as follows:

$$\max \sum_{i=1}^{n} [C_{nB}(n^*) - C_{nB}(n_i)]$$

$$= (n-1) * [1-0] + 1 * [1-1] = (n-1)$$

$$\text{with } C_B(n^*) = \frac{n^2 - 3n + 2}{2}, C_B(n_i) = 0$$

Since the observed network is always related to the maximum possible network centralization, as C_D and C_C, C_B ranges in the interval [0, 1]. A low value implies that – according to the betweenness-based measure – a decentral network is present and a high value implies that a rather centralized network is present. In example network A (see p. 81, Figure 19) C_B=1, in network B, C_B=5/12.

Betweenness centrality measures the control and profit options which an actor has due to its structural position in a network (see JANSEN 2006: 135). One positive aspect is that the (normalized) betweenness centrality and betweenness centralization are always computable. According to FREEMAN (1979: 222) the difficulty with betweenness centrality is that when there are several geodesics connecting a pair of points a point falling on some but not all of the geodesics connecting a pair of others has a more limited potential for control in comparison to a network where there is only one geodesic connecting each pair of points. Also, when being interested in an actor's communication activity or its dependence on others, this group of measures seems rather inappropriate to employ.

5.3.6 Results of the closeness- and betweenness-based measures

In section 3.5.2, the focus is on aspects of centralization in the networks of technological overlap: Section 3.5.2.1 serves to detect centrality in the networks of technological overlap, which is assessed using the degree-based measurement concept. Section 3.5.2.2 then follows with implications which can be derived from the observed centrality patterns. For reasons of completion and comparability, the findings for the other two measurement concepts, the closeness- and betweenness-based measurement concept, are presented in the following two sections. Since in the networks of technological overlap, these measurement concepts do not come along with a natural interpretation, section 5.3.6.1 and section 5.3.6.2 merely focus on the presentation of the results.

5.3.6.1 Results of the closeness-based measures

Using degree-based measures as an indication of centrality and centralization yields a local view as only direct technological ties between companies are observed. The measures normalized closeness centrality and closeness centralization also consider the "technological" kinship, i.e. indirect ties between actors. As mentioned in section 5.3.5.1, the computation of these measures is problematic in case of unconnected network parts. To obtain the normalized closeness centralities of the vertices, the results need to be weighted. However, the measure closeness centralization of these networks cannot be computed. Since in this book, in cohort 1, 2 and 4 parts of the network are unconnected (see p. 88, Table 25), the respective normalized closeness centralities are weighted while in the third and fifth cohort, the normalized closeness centralities are not weighted. Consequently, the results presented below need to be treated with care as the normalized closeness centralities of cohort 1, 2 and 4 can only be of limited comparability. Furthermore, in these cohorts, "closeness centralization" cannot be computed. Table 36 (below) depicts the normalized closeness centrality and closeness centralization of all five cohorts.

Table 36: Normalized closeness centrality and closeness centralization of the networks of technological overlap[133].

		Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5
Distribution of the normalized closeness centralities	**(0.000, 0.200]**	7	2	0	2	0
	(0.200, 0.400]	6	0	1	0	3
	(0.400, 0.600]	5	33	41	66	159
	(0.600, 0.800]	0	6	27	53	63
	(0.800, 1.000]	0	0	3	4	6
AM		0.263	0.526	0.603	0.592	0.580
SD		0.172	0.135	0.093	0.112	0.079
Interval […]		0.000, 0.556	0.049, 0.786	0.370, 0.947	0.016, 0.903	0.328, 0.888
C_C[134]		-	-	0.702	-	0.619

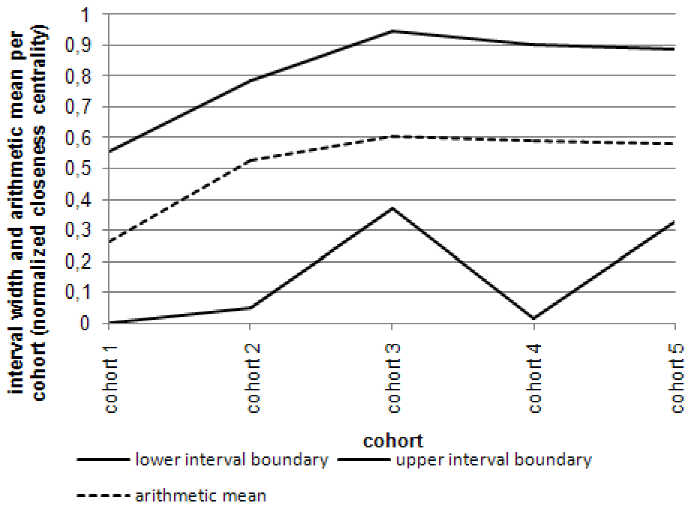
Source: O*wn computations based on dataset 2.*

[133] Multiple lines and loops are removed.
[134] Closeness centralization.

Following Table 36, with the exception of cohort 1, the *distribution* to the classes shows that the majority of companies have a *medium normalized closeness centrality* ranging between 0.400 and 0.600. Again except for the first cohort, the second largest class in every cohort is the one depicting companies with a normalized closeness centrality between 0.600 and 0.800. Few companies are technologically closer linked to other companies, i.e. more central, and few less. Consequently, in three of the five cohorts the *arithmetic mean* lies in the middle interval of (0.400; 0.600]; in case of cohort 3, the arithmetic mean even lies in the interval (0.600; 0.800]. In cohort 1, the arithmetic mean is considerably lower. In general, the higher mean normalized closeness centrality in cohort 2, 3, 4 and 5 could be a sign of closer technological kinship over time. Additionally, with the exception of the fourth cohort, the *standard deviation* depicts a decreasing trend: in the course of time, the spread from the arithmetic mean becomes smaller. This is not astonishing as in the *first* cohort, the companies have very different normalized closeness centrality values while in the fifth cohort for example, the majority of companies have a normalized closeness centrality ranging between 0.400 and 0.600.

To draw a little sharper picture of centrality, the intervals are useful to examine. On the basis of Table 36 (p. 161), the course of the upper and lower boundary as well as the arithmetic mean is shown in Figure 49 (below).

Figure 49: Interval width and arithmetic mean in all cohorts, normalized closeness centrality.



Source: Own computations based on dataset 2.

The lower boundaries of the *intervals* in which the normalized closeness centralities of the companies range vary between 0.000 in the first and 0.370 in the third cohort. In the first cohort, there are companies which are technologically unrelated (isolated); consequently, the respective actors have a normalized closeness centrality of 0.000. In the second and fourth cohort the network consists of two components and the smaller of the two components comprises only two actors. Due to the fact that the networks contain components, the respective normalized closeness centrality values of the actors are weighted. This should be kept in mind when interpreting the results. In turn, the networks in cohort 3 and 5 are not split into components, so the respective results can be compared directly. In these cohorts, the lower boundary is comparably high. Concerning the course of the upper boundary, it is striking that in the first two cohorts the upper boundary is rather low; in the remaining cohorts the upper boundary of the intervals is around 0.900. While in earlier times there do not seem to be companies which are technologically either highly directly or indirectly related to others, this seems to be the case in later times.

The course of the lower and upper boundary of the intervals shows that in cohort 1, 2, 3 and 5, the gap between companies with a low and those with a high normalized closeness centrality is alike. The network of cohort 4 seems to be the most centralized one.

Also the analysis of technological kinship shows that – at least in the network of technological overlap – large chemical companies seem to dominate the corporate landscape. In four of five cohorts, the most central actor is a large chemical company. Amongst these, one company reaches the following normalized closeness centrality values: cohort 1. 0.556; 2. 0.738; 3. 0.866; 4. 0.903; 5. 0.888. In other words, in three of five cohorts, this company determines the upper boundary of the respective interval; according to the closeness-based measures, it is the most central actor in the network of technological overlap. In cohort 2 and in cohort 3, two further companies determine the upper boundary.

As mentioned, due to the presence of unconnected parts, in the network the measure *“closeness centralization”* cannot be computed for cohort 1, 2 and 4. For the remaining two cohorts, the measure closeness centralization suggests a slightly higher tendency of centralization of the network in the third than in the fifth cohort.

Altogether, according to the closeness-based measurement concept, core and peripheral actors may be detected in the cohorts, but the overall network centralization cannot be assessed due to the presence of unconnected network parts. Therefore, from the view of the closeness-based measurement concept (which is not in focus in this book), hypothesis 5 cannot be assessed.

5.3.6.2 Results of the betweenness-based measures

The degree- and closeness-based measures are distance related measures. While the interpretation of these two measures comes more or less natural, the case is more complex in terms of the measure betweenness centrality/ centralization. Theoretically, the normalized *betweenness centrality/ centralization* may give insight into the role of a company in the sense of a company's position as an intermediary in the network. However, defining an intermediary in the networks of technological overlap is difficult to do because connections in these networks cannot be equated with actual connections, for example in the sense of cooperations.

Furthermore, as mentioned in section 5.3.5.2, the measure normalized betweenness centrality comes along with the difficulty of not adequately reflecting the potential for control in networks with more than one geodesic connecting a pair of points. However, this is likely to frequently occur in the networks presented in this book. Due to both reasons, the results presented in the following Table 37 (below) need to be treated with care.

Table 37: Normalized betweenness centrality and betweenness centralization of the networks of technological overlap[135].

		Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5
Distribution of the norm. betweenness centralities	**(0.000, 0.050]**	16	37	69	123	231
	(0.050, 0.100]	1	1	2	2	0
	(0.100, 0.150]	1	1	1	0	0
	(0.150, 0.200]	0	2	0	0	0
	(0.200, 1.000]	0	0	0	0	0
AM		0.014	0.018	0.010	0.005	0.003
SD		0.030	0.039	0.021	0.012	0.007
Interval […]		0.000, 0.114	0.000, 0.155	0.000, 0.126	0.000, 0.092	0.000, 0.047
C_B[136]		0.106	0.141	0.118	0.088	0.044

Source: Own computations based on dataset 2.

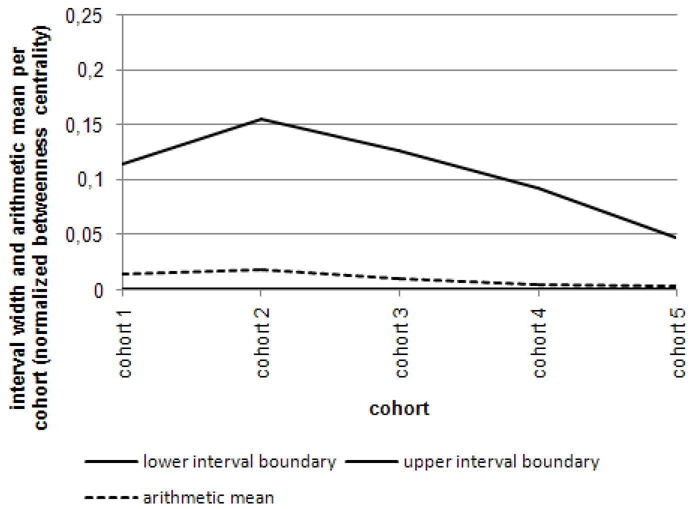
The *distribution* to the classes and cohorts shows that the majority of companies do have a normalized betweenness centrality of below 0.050. Few companies reach a higher normalized betweenness centrality, i.e. are more central. The *arithmetic mean* follows accordingly: It ranges between 0.003 in the fifth and

[135] Multiple lines and loops are removed.
[136] Betweenness centralization.

0.018 in the second cohort. The *standard deviation* reveals that (with the exception of cohort 2) the spread from the arithmetic mean becomes smaller over time.

A more profound picture of centrality can be retrieved from the intervals, in which the normalized betweenness centrality ranges. Based on the computations (see Table 37) the course of the lower and upper interval boundary as well as the arithmetic mean are shown in Figure 50 (below).

Figure 50: Interval width and arithmetic mean in all cohorts, normalized betweenness centrality.



Source: *Own computations based on dataset 2.*

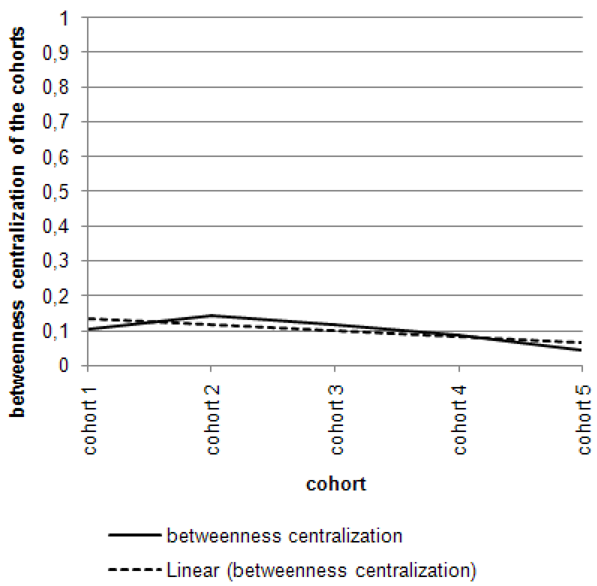
In all five cohorts, the lower *interval* boundary is equal to zero implying that in each cohort there is at least one company which is not important as a technological intermediary in the network of technological overlap. Unsurprisingly, the upper interval boundaries are also relatively low in each cohort. No actor is as important as to be the only technological link to the remaining companies. The removal of one company does not lead to the "collapse" of the whole network. Since only 35 technologies are considered, in every cohort, parts of the networks always remain connected even if one company is artificially removed.

The course of the lower and upper boundary of the intervals shows that from cohort 2 on, the gap between companies with a low and those with a high normalized betweenness centrality strongly decreases. The network in cohort 2 seems to be the most centralized one according to the measure of betweenness.

Other than the degree and closeness measures, the measure betweenness centrality partially suggests different companies to be central. One chemical company for example reaches the following normalized betweenness centrality values: cohort 1. 0.114; 2. 0.126; 3. 0.085; 4. 0.067; 5. 0.044. Obviously, in the first cohort this actor is the most central actor. A closer look at the underlying data furthermore reveals that – when considering the measure normalized betweenness centrality – the company always is amongst the top three central actors. Else, in two of the remaining four cohorts, the most central actors are other large chemical companies. Finally, in the two remaining cohorts the most central actor is an actor with its main emphasis on fields other than chemistry.

Finally, the company specific results or rather *the normalized betweenness centrality* can be held against the betweenness centralization of the entire network to gain an overall impression of network centrality. For reasons of illustration, in addition to Table 37 (p. 172), the betweenness centralization of the networks is shown in Figure 51 (below).

Figure 51: Development of betweenness centralization.



Source: Own computations based on dataset 2.

The measure *"betweenness centralization"* confirms the impression of rather decentralized networks. It is considerably low in all cohorts varying between 0.044 in the fifth cohort and 0.141 in the second. The measure betweenness centralization signals a decreasing tendency of centralization from cohort 2 on. The existent central actors are never the only technological links between companies and over time, in this respect, their relevance decreases. Since only 35 technologies are considered parts of the networks always remain intact even if one company is artificially removed.

Altogether, according to the betweenness-based measurement concept, core and peripheral actors cannot really be detected in the cohorts. Also, the overall network centralization shows a decreasing tendency. From the view of the closeness-based measurement concept (which is not in focus in this book) hypothesis 5 has to be rejected.

6 List of Abbreviations

AFT	Accelerated failure time
AIC	Akaike information criterion
AM	Arithmetic mean
BfR	Bundesinstitut für Risikobewertung
BMBF	Bundesministerium für Bildung und Forschung (engl.: Federal Ministry of Education and Research
BMWi	Bundesministerium für Wirtschaft und Technologie (engl.: Federal Ministry of Economics and Technology)
DPMA	Deutsches Patent- und Markenamt (engl. German Patent and Trade Mark Office)
EPO	European Patent Office
EPV	Events per variable
IPC	International Patent Classification
LP	Leistungsplansystematik
NACE	Nomenclature générale des activités économiques
PH	Proportional hazard
R&D	Research and Development
SCI	Science Citation Index
SD	Standard deviation
VC	Venture capital
WIPO	World Intellectual Property Organization
WZ	Klassifikation der Wirtschaftszweige (engl. German Classification of Economic Activities).

7 FIGURES AND TABLES

7.1 Figures

Figure 1: General development trend and relatedness to nanotechnology. 4
Figure 2: Distribution of NACE-codes, in percent. 21
Figure 3: Distribution per Federal State, in percent.. 23
Figure 4: Distribution by size following the classification of the EUROPÄISCHE GEMEINSCHAFTEN (2006: 14), in percent. 24
Figure 5: Foundry years. 26
Figure 6: Distribution of mergers and acquisitions per company size................ 28
Figure 7: Distribution of market exits per company size................................... 29
Figure 8: Kaplan-Meier survival curve. 43
Figure 9: Kaplan-Meier curves for pre-entry experience. 44
Figure 10: Kaplan-Meier curves for post-entry experience................................ 46
Figure 11: Kaplan-Meier curves for technological know-how. 47
Figure 12: Smoothed hazard. 50
Figure 13: Cox-Snell residuals of model 5 and model 6.................................... 57
Figure 14: Post-entry experience, Martingale residuals for model 5. 58
Figure 15: Deviance residuals for model 5 and model 6. 59
Figure 16: Distribution of the types of applicants. 73
Figure 17: Timely development of patent applications. 75
Figure 18: Distribution of patent applications to the 35 fields of technology on the basis of the WIPO IPC-Technology Concordance Table. 77
Figure 19: Exemplary networks. 81
Figure 20: Network of technological overlap, cohort 1. 91
Figure 21: Network of technological overlap, cohort 2. 92
Figure 22: Network of technological overlap, cohort 3. 93
Figure 23: Network of technological overlap, cohort 4. 94
Figure 24: Network of technological overlap, cohort 5. 95
Figure 25: Interval width and arithmetic mean in all cohorts, normalized degree centrality.. 98
Figure 26: Development of degree centralization. 99
Figure 27: Permanent and exiting companies, cohort 1. 103
Figure 28: Permanent and exiting companies, cohort 2. 104
Figure 29: Permanent and exiting companies, cohort 3. 105
Figure 30: Permanent and exiting companies, cohort 4. 106
Figure 31: Sensitivity analysis for cohort 1. 108
Figure 32: Sensitivity analysis for cohort 2. 109
Figure 33: Sensitivity analysis for cohort 3. 110
Figure 34: Sensitivity analysis for cohort 4. 111
Figure 35: Development of top five chosen technologies................................ 113
Figure 36: Technological profile of cohort 1, in percent. 114
Figure 37: Technological profile of cohort 2, in percent. 114
Figure 38: Technological profile of cohort 3, in percent. 115
Figure 39: Technological profile of cohort 4, in percent. 115
Figure 40: Expenditures for nanotechnology within the scope of diverse BMBF key issues. 132
Figure 41: Occurence of ties in dataset 1. 143
Figure 42: Hazard rates of the fitted fully parametric models. 146
Figure 43: Kaplan-Meier curves for the company size. 147
Figure 44: Exemplary patent application at the DPMA.................................... 148
Figure 45: Cooperation in the network of technological overlap, cohort 3. 159

Figure 46: Cooperation in the network of technological overlap, cohort 4. 160
Figure 47: Cooperation in the network of technological overlap, cohort 5. 161
Figure 48: WIPO IPC-Technology Concordance Table. 163
Figure 49: Interval width and arithmetic mean in all cohorts, normalized closeness centrality. 170
Figure 50: Interval width and arithmetic mean in all cohorts, normalized betweenness centrality. 173
Figure 51: Development of betweenness centralization. 174

7.2 Tables

Table 1: Förderkatalog, all LP-numbers and labels. 16
Table 2: Förderkatalog, main LP-numbers and labels. 17
Table 3: Nanoproducts, main and subgroups. 19
Table 4: Top ten chosen NACE-codes in the sample, in percent. 22
Table 5: Company statuses in the database, in percent. 27
Table 6: Computation of duration times, censoring. 30
Table 7: Types of entry in the database, in percent. 31
Table 8: Determination of diversifiers. 33
Table 9: Distribution over entry cohorts. 34
Table 10: Distribution of patent applications, in percent. 35
Table 11: Overview on PH- and AFT-models. 37
Table 12: Examined covariates possibly influencing the survival of companies. 41
Table 13: Further statistics of the dataset. 42
Table 14: Pre-entry experience, extended mean. 45
Table 15: Post-entry experience, extended mean. 47
Table 16: Technological know-how, extended mean. 48
Table 17: Model configuration. 51
Table 18: Cox regression (efron method for ties). 53
Table 19: Stratified Cox regression (efron method for ties). 53
Table 20: Interaction effects (Cox regression, efron method for ties). 55
Table 21: Link test for model 5 and model 6. 59
Table 22: Search strategy for the identification of applicant types. 72
Table 23: Distribution of companies over the cohorts. 74
Table 24: IPC-sections. 76
Table 25: Descriptive statistics of the networks of technological overlap. 88
Table 26: Normalized degree centrality and degree centralization of the networks of technological overlap. 97
Table 27: Contingency table. 107
Table 28: Mean normalized degree centrality of permanent and exiting companies, cohorts 1-4. 112
Table 29: Keyword list. 133
Table 30: Search strategy for patent applications in nanotechnology. 134
Table 31: NACE-keys. 139
Table 32: Cox regression (efron method for ties), Schoenfeld residuals. 144
Table 33: Results of the estat phtest. 144
Table 34: AIC for PH-models. 145
Table 35: AIC for AFT-models. 145
Table 36: Normalized closeness centrality and closeness centralization of the networks of technological overlap. 169
Table 37: Normalized betweenness centrality and betweenness centralization of the networks of technological overlap. 172

8 BIBLIOGRAPHY

Agarwal, R.; Echambadi, R.; Franco, A. M.; Sarkar, M. B. (2004): Knowledge Transfer Through Inheritance: Spin-out Generation, Development, and Survival, in: Academy of Management Journal (2004), Vol. 47, No. 4: 501-522.

Ahuja, G. (2000): Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study, in: Administrative Science Quarterly (2000) 45: 425-455.

Autant-Bernard, C.; Billand, P.; Bravard, C.; Massard, N. (2007): Network Effects in R&D Partnership Evidence from the European Collaborations in Micro and Nanotechnologies, http://hal-ujm.ccsd.cnrs.fr/docs/00/13/72/38/PDF/DIME-07.pdf, 28 October 2009.

Baumgartner, W.; Jäckli, B.; Schmithüsen, B.; Weber, F. (2003): Nanotechnologie in der Medizin, Studie des Zentrums für Technologiefolgen-Abschätzung, http://www.ta-swiss.ch/a/biot_nano/2003_47_nanotechnologie_d.pdf, 21 October 2009.

Bhushan, B. (Ed.) (2005): Springer handbook of nanotechnology (El. Resource), Norwich, NY: Knovel.

BioM AG Munich BioTech Development (2005): Die BioTech-Region München – Erfolg schafft neue Zuversicht, http://www.bio-m.org/_resources/dynamic/hauptbereich/bio_m_publikationen/report_2005_deutsch.pdf, 10 August 2009.

BioRiver – Life Science im Rheinland e.V. (2007): BioRiver Report 2007, http://www.bioriver.de/fileadmin/downloads/BioRiver_Report_Deu.pdf, 10 August 2009.

BioTOP Berlin-Brandenburg (2008): Biotech-Report 2007·2008, Ausgabe 34 Mai 2008, http://www.biotop.de/data/files/biotopics/Branchenreport2008-D.pdf, 10 August 2009.

Boeing, N. (2008): Nanotechnologien – Konzepte, Perspektiven, Risiken, http://www.verbraucherbildung.de/projekt01/media/pdf/FB_Nanotechnik_Boeing_112008.pdf, 1 April 2010.

Boisseau, P. (2007): Nano2Life, European network of excellence in nanobiotechnology, in: EuroNanoForum 2007 – Nanotechnology in Industrial Applications, European and International Forum of Nanotechnology, Düsseldorf, 19-21 June 2007: 213.

Bradburn, M. J.; Clark, T. G.; Love, S. B.; Altman, D. G. (2003A): Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods, in: British Journal of Cancer (2003) 89: 431-436.

Bradburn, M. J.; Clark, T. G.; Love, S. B.; Altman, D. G. (2003B): Survival Analysis Part III: Multivariate data analysis – choosing a model and assessing its adequacy and fit, in: British Journal of Cancer (2003) 89: 605-611.

Buenstorf, G. (2005): Evolution on the Shoulders of Giants: Entrepreneurship and Firm Survival in the German Laser Industry, in: Papers on Economics & Evolution: 1-35.

Buenstorf, G. (2007): Evolution on the Shoulders of Giants: Entrepreneurship and Firm Survival in the German Laser Industry, in: Rev Ind Organ (2007) 30: 179-202.

Bundesinstitut für Risikobewertung (2006): BfR-Verbraucherkonferenz zur Nanotechnologie in Lebensmitteln, Kosmetika und Textilien, URL: http://www.bfr.bund.de/cm/220/verbrauchervotum_zur_nanotechnologie.pdf, 18 June 2009.

Bundesinstitut für Risikobewertung: Nanopartikel waren nicht die Ursache für Gesundheitsprobleme durch Versiegelungssprays!, http://www.bfr.bund.de/cd/7839, 18 June 2009.

Bundesministerium für Bildung und Forschung (Ed.) (2004): Nanotechnology Conquers Markets, German Innovation Initiative for Nanotechnology, http://www.bmbf.de/pub/nanotechnology_conquers_markets.pdf, 21 October 2009.

Bundesministerium für Bildung und Forschung (Ed.) (2006A): Nanotechnologie, Innovationen für die Welt von morgen, http://www.bmbf.de/pub/nanotechnologie_inno_fuer_die_welt_v_morgen.pdf, 21 October 2009.

Bundesministerium für Bildung und Forschung (Ed.) (2006B): Nano-Initiative – Aktionsplan 2010, http://www.bmbf.de/pub/nano_initiative_aktionsplan_2010.pdf, 21 October 2009.

Bundesministerium für Bildung und Forschung/ Bundesministerium für Wirtschaft und Technologie: Förderkatalog, http://oas2.ip.kp.dlr.de/foekat/foekat/foekat, 16 January 2007.

Burr, W.; Grupp, H.; Funken-Vrohlings, M. (2009): Regulierung und Produkthaftung in einem jungen Technologiefeld, am Beispiel der Nanotechnologie, in:

Nanotechnologie, Grundlagen, Anwendungen, Risiken, Regulierung, Berlin: De Gruyter Recht: 249-275.

Cantner, U.; Dreßler, K.; Krüger, J. J. (2005): Knowledge and Creative Destruction over the Industry Life Cycle – The Case of the German Automobile Industry, in: Jenaer Schriften zur Wirtschaftswissenschaft: 1-23.

Cantner, U.; Dreßler, K.; Krüger, J. J. (2006): Firm survival in the German automobile industry, in: Empirica (2006) 33: 49-60.

Cantner, U. and Graf, H. (2006): The network of innovators in Jena: An application of social network analysis, in: Research Policy (2006) 35: 463-480.

Charisius, H.: Falsche Furcht, http://www.heise.de/tr/Falsche-Furcht--/blog/artikel/71929, 18 June 2009.

Clark, T. G.; Bradburn, M. J.; Love; S. B.; Altman, D. G. (2003): Survival Analysis Part I: Basic concepts and first analyses, in: British Journal of Cancer (2003) 89: 232-238.

Cleves, M.; Gould, W. W.; Gutierrez, R. G.; Marchenko, Y. U. (2008): An Introduction to Survival Analysis Using Stata, 2nd edition, Texas: Stata Press Publication.

DEPATISnet: http://depatisnet.dpma.de, 5 January 2011.

Deutsches Patent- und Markenamt: Glossar, http://www.deutsches-patentamt.de/service/glossar/n_r/index.html#a1, 10 December 2010.

EPO: History, http://www.epo.org/about-us/office/history.html, 21 December 2010.

EPO Worldwide Statistical Patent Database version October 2007 (PATSTAT version 10/ 2007), published by the Patent Statistic Task Force.

Europäische Gemeinschaften (2006): Die neue KMU-Definition – Benutzerhandbuch und Mustererklärung, http://ec.europa.eu/enterprise/enterprise_policy/sme_definition/sme_user_guide_de.pdf, 14 December 2007.

Feynman, R. P. (1992): There's Plenty of Room at the Bottom, reprinted with permission from Journal of Microelectromechanical Systems, Vol. 1, No. i: 60-66, March 1992, Originally published in Miniaturization Horace D. Gilbert, Ed.

Freeman, L. C. (1979): Centrality in Social Networks Conceptual Clarification, in: Social Networks (1978/ 1979) 1: 215-239.

Frietsch, R.; Köhler, F.; Blind, K. (2008): Weltmarktpatente – Strukturen und deren Veränderungen, Studien zum deutschen Innovationssystem Nr. 7-2008.

Führ, M.; Hermann, A.; Merenyi, S.; Moch, K.; Möller, M.; Kleihauer, S.; Steffensen, B. (2006): Rechtsgutachten Nano-Technologien – ReNaTe, Bestehender Rechtsrahmen, Regulierungsbedarf sowie Regulierungsmöglichkeiten auf europäischer und nationaler Ebene, http://oeko.de/oekodoc/334/2006-022-de.pdf, 22 October 2009.

Gesellschaft für Bioanalytik Münster e. V. (2004): Münster biotech-region companies and competence, http://www.bioanalytik-muenster.de/conpresso/_data/Biotechregion_3-1.pdf, 10 August 2009.

GENIOS German Business Information (GENIOS): http://www.genios.de/, 9 January 2011.

Gobrecht, J. (2006): Rückblick auf die Entstehungsgeschichte der Nanotechnologie, in: Highlights aus der Nano-Welt Eine Schlüsseltechnologie verändert unsere Gesellschaft, Freiburg: HERDER spektrum: 12-30.

Granovetter, M. (1983): The Strength of Weak Ties: A Network Theory Revisited, in: Sociological Theory (1983) 1: 201-233.

Grupp, H. (1997): Messung und Erklärung des Technischen Wandels Grundzüge einer empirischen Innovationsökonomik, Berlin: Springer Verlag.

Heinze, T. (2006): Die Kopplung von Wissenschaft und Wirtschaft Das Beispiel der Nanotechnologie, Frankfurt: Campus Verlag.

Hessen-Nanotechnologie: KOMPETENZMATRIX UNTERNEHMEN 1, http://www.hessen-nanotech.de/mm/02_Kompetenzmatrix.pdf, 6 November 2008.

Heubach, D.; Beucker, S.; Lang-Koetz, C. (2005): Einsatz von Nanotechnologie in der hessischen Umwelttechnologie, Innovationspotentiale für Unternehmen, http://www.hessen-nanotech.de/mm/Nanoinut_Endversion_Internet.pdf, 21 October 2009.

Holister, P. (2002): Nanotech The Tiny Revolution, http://www.nanotech-now.com/CMP-reports/NOR_White_Paper-July2002.pdf, 22 October 2009.

Hoppenstedt: Firmenadressen für Direktmarketing, http://www.hoppenstedt-adressen.de/?gclid=CKqQ_5Kur6YCFUYifAodmBQfjg, 10 January 2011.

Hoppenstedt (2009): Hoppenstedt Firmendatenbank Branchen Nomenklatur WZ 2008 mit Schlüssel und Codierungen.

Hoppenstedt Firmendatenbank (Hoppenstedt), run by Hoppenstedt.

Horvath, M.; Schivardi, F.; Woywode, M. (2000): On Industry Life-Cycles: Delay, Entry, and Shakeout in Beer Brewing, http://digidownload.libero.it/fschivardi/images/brijo20.pdf, 14 August 2009.

Hullmann, A. (2001): Internationaler Wissenstransfer und technischer Wandel – Bedeutung, Einflussfaktoren und Ausblick am Beispiel der Nanotechnologie in Deutschland, Heidelberg: Physica-Verlag.

Hullmann, A. and Fryček, R. (2007): IPR in Nanotechnology - lessons from experiences worldwide Proceedings, Workshop in the frame of the INC3 conference, jointly organised by the European Patent Office and the European Commission, DG Research, 16 April 2007, Brussels, Belgium.

Jansen, D. (2006): Einführung in die Netzwerkanalyse Grundlagen, Methoden, Forschungsbeispiele, 3. Auflage, Wiesbaden: VS Verlag für Sozialwissenschaften/GWV Fachverlage GmbH.

Klepper, S. (2002a): The capabilities of new firms and the evolution of the US automobile industry, in: Industrial and Corporate Change (2002), Vol. 11, No. 4: 645-666.

Klepper, S. (2002b): Firm Survival and the Evolution of Oligopoly, in: The RAND Journal of Economics (2002), Vol. 33, No. 1: 37-61.

Klepper, S. and Simons, K. L. (1999): Industry Shakeouts and Technological Change, http://webu2.upmf-grenoble.fr/iepe/textes/KlepperSimons.PDF, 29 November 2006.

Klepper, S. and Thompson, P. (2006): Intra-Industry Spinoffs, http://www.fiu.edu/orgs/economics/wp2006/06-05.pdf, 31 January 2007.

Kühling, W. and Horn, H. (2007): Für einen verantwortungsvollen Umgang mit der Nanotechnologie, Eine erste Diskussionsgrundlage am Beispiel der Nanopartikel, http://www.ulrike-hoefken.de/cms/default/dokbin/269/269806.position_bund_nano.pdf, 16 June 2009.

Landesentwicklungsgesellschaft (LEG) Thüringen mbH Abteilung Akquisition und internationale Kontakte (2004): Biotechnologie in Thüringen, http://www.leg-thueringen.de/fileadmin/www/pdfs/DE/publikationen/factsheet_bio.pdf, 10 August 2009.

LexisNexis, run by LexisNexis Deutschland GmbH.

Morrison, M. (2007): The outcomes and future of Nanoforum, European nanotechnology network, in: EuroNanoForum 2007 – Nanotechnology in Industrial Applications, European and International Forum of Nanotechnology, Düsseldorf, 19-21 June 2007: 206.

Nanoproducts: http://www.nanoproducts.de/, 9 March 2007 and http://www.nanoproducts.de/index.php?mp=info&file=nanotechnologie, 21 October 2009.

Nanotechnologie in Hessen: Nanotechnologie-Unternehmen, http://www.nanoportal-hessen.de/brancheninfo/unternehmen/, 9 March 2007.

Nanotechnology Research Foundation, http://nanotechnologyresearchfoundation.org/nanohistory.html, 16 March 2009.

Nano-Handel.de: Nanotechnologie für Handel, Gewerbe & Industrie, http://nano-handel.de/handel-gewerbe-industrie.php?shop=d94d09f05797f3373b00a925db57181c, 10 September 2011.

NetzwerkTotal Das Netzwerk Portal: http://www.netzwerktotal.de/netzwerktopologie.htm, 19 January 2010.

Neuhäusler, P. (2008): Patente in Europa und den USA – Veränderungen ab 1991 aufgezeigt an Gesamtzahlen und dem Technologiefeld des Ubiquitous Computing, in: Fraunhofer ISI Discussion Papers *Innovation Systems and Policy Analysis*, No. 14: 1-34.

Noack, M. (2008): Cox-Regression, http://www.uni-due.de/imperia/md/content/soziologie/stein/beamer-ereignis2nopause.pdf, 14 August 2009.

De Nooy, W.; Mrvar, A.; Batagelj, V. (2005): Exploratory Social Network Analysis with Pajek, New York: Cambridge University Press.

Noyons, E. C. M.; Buter, R. K.; van Raan, A. F. J.; Schmoch, U.; Heinze, T.; Hinze, S.; Rangnow, R. (2003): Mapping Excellence in Science and Technology across Europe Nanoscience and Nanotechnology, Draft report of project EC-PPN CT-2002-0001 to the European Commission: 1-113.

Oberdörster, E. (2004): Manufactured Nanomaterials (Fullerenes, C_{60}) Induce Oxidative Stress in the Brain of Juvenile Largemouth Bass, in: Environmental Health Perspectives (2004) 112, No. 10: 1058-1062.

Oberdörster, E.; Zhu, S.; Blickley, T. M.; McClellan-Green, P.; Haasch, M. L. (2005): Ecotoxicology of carbon-based engineered nanoparticles: Effects of fullerene (C_{60}) on aquatic organisms, in: Carbon (2006) 44: 1112-1120.

Ott, I. and Papilloud,C. (2007): Converging Institutions: Shaping Relationships Between Nanotechnologies, Economy, and Society, in: Bulletin of Science, Technology & Society Vol. 27, No. 6, December 2007, 455-466.

Peduzzi, P.; Concato, J.; Feinstein, A. R.; Holford, T. R. (1995): Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates, in: J Clin Epidemiol Vol. 48, No. 12: 1503-1510.

Powell, W. W.; Koput, K. W.; Smith-Doerr, L. (1996): Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology, in: Administrative Science Quarterly (1996) 41: 116-145.

Scheidegger, N. (2008): Die Wirkung struktureller Löcher auf den Karriereerfolg im Management: eine kontingente Betrachtung, in: Netzwerkanalyse und Netzwerktheorie Ein neues Paradigma in den Sozialwissenschaften, C. Stegbauer (Ed.), Wiesbaden: VS Verlag für Sozialwissenschaften|GWV Fachverlage GmbH: 503-516.

Scheu, M; Veefkind, V.; Verbandt, Y.; Galan, E. M.; Absalom, R.; Förster, W. (2006): Mapping nanotechnology patents: The EPO approach, in: World Patent Information (2006) 28: 204-211.

Schmoch, U. (2008): Concept of a Technology Classification for Country Comparisons Final Report to the World Intellectual Property Organization (WIPO), http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/pdf/wipo_ipc_technology.pdf, 10 October 2009.

Stanley, E. A. (2006): Social Networks and Mathematical Modeling, in: Connections (2006) 27: 39-45.

StataCorp LP (2007): Stata Statistical Software: Release 10 Survival Analysis and Epidemiological Tables, College Station, TX: Stata Press.

Stöber, S. and Türk, V. (2006): Hoffnungsträger Nanotechnologie, Ein Dialog über gesellschaftliche und ethische Fragen einer neuen Technologie, in: zfwu – Zeitschrift für Wirtschafts- und Unternehmensethik (2006) 7/2: 277-291.

Sweeney, K.: Duration Analysis In Stata, psweb.sbs.ohio-state.edu/faculty/jbox/.../statanotes.pdf, June 4, 2011.

Sydow, J. (1992): Strategische Netzwerke Evolution und Organisation, Wiesbaden: Gabler.

Technologiepark Heidelberg (2004): Companies in the Field of Biotechnology BioRegion Rhein-Neckar-Dreieck, http://www.citebiotech.com/Docs/fra/Doc%20partenaires/Heidelberg%202004-09-29.pdf, 10 August 2009.

Teece, D. J. (1992): Competition, cooperation, and innovation, in: Journal of Economic Behavior and Organization (1992) 18: 1-25.

Thompson, P. (2003): Selection and Firm Survival – Evidence from the Shipbuilding Industry, 1825-1914, http://www.hbs.edu/units/tom/seminars03-04/pthompson.pdf, 29 January 2007.

UCLA Academic Technology Services: Stata Web Books Regression with Stata Chapter 2 – Regression Diagnostics, http://www.ats.ucla.edu/stat/stata/webbooks/reg/chapter2/statareg2.htm, 27 July 2009.

Umweltbundesamt (2006): Nanotechnik: Chancen und Risiken für Mensch und Umwelt Hintergrundpapier August 2006, http://www.umweltbundesamt.de/uba-info-presse/hintergrund/nanotechnik.pdf, 17 June 2009.

VDI-Technologiezentrum: Nanomap, http://www.nano-map.de/, 7 August 2007.

Wasserman, S. and Faust, K. (1994): SOCIAL NETWORK ANALYSIS: METHODS AND APPLICATIONS, Cambridge: Cambridge University Press.

World Intellectual Property Organization (2009): INTERNATIONAL PATENT CLASSIFICATION (Version 2009) GUIDE http://www.wipo.int/export/sites/www/classifications/ipc/en/guide/guide_ipc_2009.pdf, 10 October 2009.

Zentrum der Gesundheit (2009): Sonnenschutz mit Nanoteilchen, http://www.zentrum-der-gesundheit.de/sonnenschutz-mit-nanoteilchen-ia.html, 17 June 2009.

Ziegler, A.; Lange, S.; Bender, R. (2004): Überlebenszeitanalyse: Die Cox-Regression – Artikel Nr. 16 der Statistik-Serie in der DMW –, in: Dtsch Med Wochenschr 2004; 129: 1-3.

Zukünftige Technologien Consulting der VDI Technologiezentrum GmbH (Ed.) (2004): Nanotechnologie als wirtschaftlicher Wachstumsmarkt Innovations- und Technikanalyse, http://www.bmbf.de/pub/nanotech_als_wachstumsmarkt.pdf, 21 October 2009.