Dimensions of Spatial Knowledge Diffusion

Technology, Network and Regional Context

Dimensies van ruimtelijke kennisverspreiding

Technologie, netwerk en regionale context

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op

dinsdag 1 september 2020 des middags te 12.45 uur

door

Marcel Bednarz

geboren op 18 februari 1990
te Bielefeld, Duitsland

Promotoren:

Prof. dr. R. A. Boschma

Prof. dr. T. Broekel

Co-promotoren:

Dr. A. Ascani

Committee:

Prof. dr. C. Autant-Bernard

Dr. P.-A. Balland

Prof. dr. C. Castaldi

Prof. dr. K. Frenken

Prof. dr. A. Varga

“More than anything else technology creates our world. It creates our wealth, our economy, our very way of being”

(Arthur, 2009: 14)

Contents

Introduction

1.1 General motivation

The European carpet of regional competitiveness is highly colorful, ranging from strong greens to deep reds, as regions vary strongly regarding their economic competitiveness and, consequently, their welfare (see Figure 1.1). The "EU Regional Competitiveness Index 2019" (RCI) published by the European Union, weaves this pattern once again (Annoni and Dijkstra, 2019). The EU assigns a competitiveness score to 268 regions, ranging from minus 1.6 to 1.08. In their index, they consider aspects such as quality of infrastructure, education of workforce and patent applications. With a score of 1.08, Stockholm leads the ranking of regional competitiveness, followed by London (1.06) and Utrecht (1.05). On the other side of the color spectrum, we find regions like the North Aegean in Greece (-1.61) and Sud-Est in Romania (-1.46).

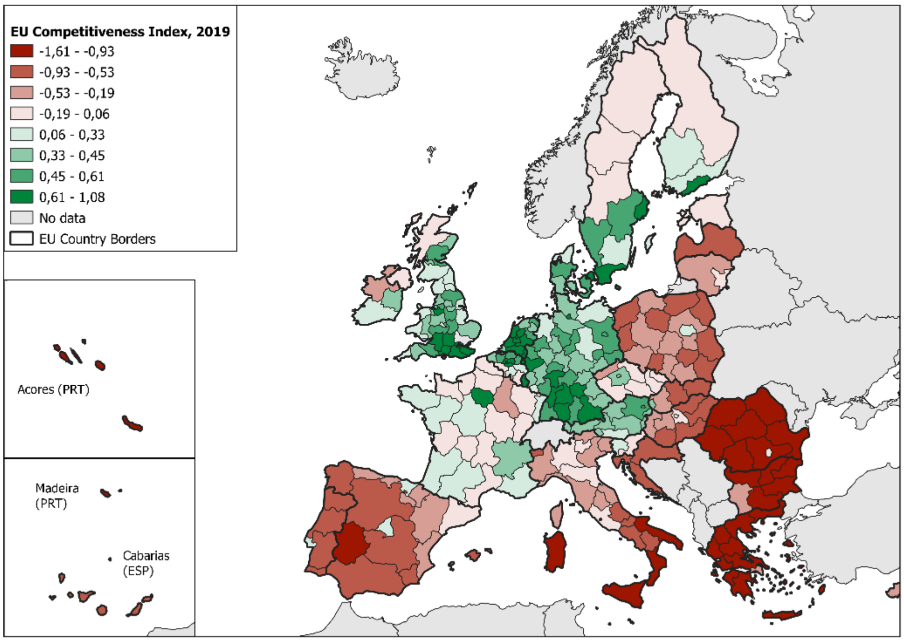

Figure 1.1: Regional Competitiveness Index 2019 for EU countries (Source: based on Annoni and Dijkstra, 2019)

In the literature, one aspect identified to define the competitiveness of regions is their capability to generate new technologies and products (Porter, 1990). Innovations enable regions to increase their relative competitiveness and generate income and wealth. However, in accordance with the picture of regional competitiveness, innovation activities concentrate in space and only a few regions are capable of regularly generating new products (Feldman, 1993;

Acs et al., 2002; Balland and Rigby, 2017). One reason for this is the cumulative nature of knowledge generation. New knowledge always builds on information and experiences people have beforehand. Thus, regions that already have less access to knowledge are less capable of producing new knowledge and fall behind. Consequently, the economic disparity between regions tends to increase (Rodríguez-Pose and Crescenzi, 2008).

Inter-regional diffusion of knowledge is one mechanism that can close such gaps. By sharing and transferring knowledge between regions, regions that did not engage in the invention of it can still utilize the knowledge in their quest for generating new technologies and products. Thereby, they are able to increase their competitiveness and eventually benefit from additional economic wealth. Like the production of knowledge, however, knowledge’s diffusion appears to be spatially sticky as well (Hägerstrand, 1952; Jaffe et al., 1993; Audretsch and Feldman, 1996; Feldman et al., 2015). Some knowledge resists diffusion and is bound to its place of origin; hence, little diffusion is observed.

But why is this so? Why do some regions adopt new knowledge much faster than other regions? Why do regions diversify into certain technologies and reject others? Can policy makers enhance the diffusion of technological knowledge in general? And if so, what tools are most appropriate for this? To answer these and related questions, a vital understanding of the mechanisms of spatial knowledge and technology diffusion is indispensable. This motivates the current thesis, which will analyze different dimensions of spatial knowledge diffusion and shed light on the adoption of technologies.

1.2 Dimensions of spatial knowledge diffusion

In his seminal work from 1890, “The laws of imitation,” translated into English in 1903, Gabriel Tarde explains the generation and diffusion of innovations as an s-shaped process: At the beginning, only few people adopt innovations. If these first adopters are satisfied with the innovation and they start using it frequently, other people will start imitating the pioneering users by trying out the innovation as well, and a re-enforcing process begins. This is the “growth phase” of an innovation, potentially leading to a broad market penetration. At the end, the process slows down as the market gets saturated and an innovation reaches its “maturity phase.” This concept has been widely used and modified in different research fields (see for example, Klepper, 1993; Rogers, 2003; Geels and Schot, 2007).

In the 1950s and 1960s, the first geographer to analyze the diffusion of innovation was the Swede Torsten Hägerstrand; he built up a computer aided simulation model with mapped data of the diffusion of cars and radios in Sweden from 1918 until 1930 (Hägerstrand, 1952; 1965).

By doing so, Hägerstrand identified a robust time lag between regions in adopting innovations. More precisely, he described three stages of the diffusion process: the "primary stage", the "diffusion stage", and the "condensing stage" (Hägerstrand, 1952, p. 16 f). In the primary stage, multiple centers of adoption rapidly appear, due to unevenly distributed information about the innovation. Originating from these centers, the information spills over into neighboring regions along the networks of social contacts. In the second phase, an increasing adoption of the innovation can be observed in these regions. New centers of adoption might arise. In general, the differences between the regions will become levelled until the "phenomenon in question is […] commonly known" (Hägerstrand, 1952, p. 17).

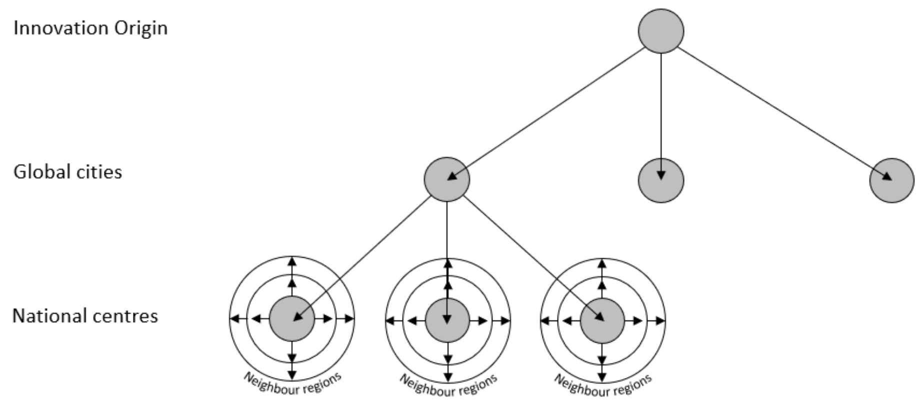

Figure 1.2: Hierarchical and wave patterns of spatial diffusion (Source: based on Kulke, 2006)

Hägerstrand (1952) identifies two spatial patterns defining these processes: first, a diffusion driven by the urban hierarchy of cities and, second, an innovation wave originating from these centers to neighboring regions (see Figure 1.2). According to his argumentation, both patterns are the results of actors' individual communication and information fields. Adopters exchange information and experiences about new innovations with potential adopters and thereby convince them to try out the innovation themselves. These exchanges happen more frequently between actors of central cities and within geographic proximity (Hägerstrand, 1967). Thus, the process of spatial diffusion is mainly determined by overcoming the ignorance about innovation through networks of private communication.

Blaut (1977) takes a rather contrasting position on this, as, for him, a diffusion model which merely concentrates on information flows "is not a simulation of the real world of geographic change" (p. 344). He criticizes Hägerstrand for emphasizing the ignorance of people as the main barrier to technology diffusion. Instead, Blaut (1977) argues that the successful acceptance of an innovation depends on the cultural context in which people live. Ormrod (1990), too, emphasizes the role of local contexts as a main driver of spatial differences in the diffusion of innovations. With regards to the increasing mass communication and, thus, the availability of information, he states that the importance of a central position within personal communication networks decreases**.** Therefore, he extends the theory of diffusion by the "concept of receptiveness." Accordingly, innovations are generated in a specific context and their successful diffusion depends on the potential customer's perception of whether the innovation is capable of providing a benefit within the customer's local context.

In more recent terms, the customer's knowledge about and his or her perception of an innovation may be affected by his or her level of proximity to the innovator. Actors living near each other have higher probabilities of meeting and interact. In this regard, geographic proximity facilitates the diffusion of knowledge between those actors (Jaffe et al., 1993; Audretsch and Feldman, 1996). However, a small physical distance may not be sufficient for a successful exchange. In this regard, Boschma (2005) identifies five dimensions of proximity: geographic, cognitive, social, institutional and organizational. Cognitive proximity, for example, affects how likely actors might make sense of new knowledge (Nooteboom et al., 2007). If an actor becomes aware of new technologies outside his field of expertise, he or she may not understand their working mechanisms and, consequently, won't use the new technologies. However, if it is related to what he or she already knows, the likelihood of successful transfers increases significantly (Nooteboom et al., 2007). Thus, in addition to the effects of regional contexts, research in Evolutionary Economic Geography (EEG) focuses on the relational level of firms and regions (Boschma, 2005).

Similar to Hägerstrand (1966), who describes the interaction of actors as networks of social contacts, EEG is also strongly interested in the structures and mechanisms of networks and how these shape the diffusion of knowledge (Glückler, 2007; Ter Wal and Boschma, 2009; Boschma and Frenken, 2010). These networks may consist of collaborating inventors who work together on one patent (Ter Wal, 2014) or organizations that participate in the same joint project (Paier and Scherngell, 2011; Broekel and Hartog, 2013a). Empirical evidence exists that the formation of these relationships is affected by proximities (Knoben and Oerlemans, 2006). For example, two actors that live in the same city are more likely to interact often and, hence, build strong relationships that may lead to frequent knowledge exchanges (Ter Wal, 2014).

In recent years a new notion was introduced to EEG, namely, complexity. For example, the production of new patents may be simple or complex, depending on the technologies used to generate this patent. Technologies are argued to be complex if they consist of numerous components and require ample information for reproduction (Simon, 1962; Kaufman, 1993).

Understanding such technologies is more difficult and time consuming (Jovanovic and Nyarko, 1995) and, thus, complex technologies tend to resist diffusion (Feldman, 1993). To create new patents under these circumstances requires collaborations between experts with deep and complementary knowledge (Balland et al., 2020).

In conclusion, generated knowledge tends to be translated into innovations in the form of new technologies. Knowledge about these technologies then diffuses through networks of personal contacts of those interacting with each other, for example, collaborating inventors. Whether these inventors then adopt the new knowledge or technology depends on their regional context, for example, the complementary capabilities and infrastructure to which they have access. This indicates three dimensions shaping the diffusion of knowledge: technology (complexity and diffusion patterns), networks (social interactions) and regional context (variations between regions in adoption and creation). The following sections will take a deep dive to each dimension and show gaps in the literature that motivate the research of this thesis.

1.2.1 Technology

The first dimension of knowledge diffusion, tackled in this thesis, is technology, as it represents the embodiment of knowledge. Novel technologies are the outcome of actors exploring and testing new combinations of knowledge and technological components (Arthur, 2009). This typically necessitates experts of different components cooperating with each other to communicate and combine their joint knowledge (Powell et al., 1996; Witt et al., 2012). However, knowledge differs in its characteristics and thus varies in its difficulty to explain. A common distinction is between codified and tacit knowledge (Polanyi, 1966; Nelson and Winter, 1982; Gertler, 2003). Codified knowledge consists of insights and experiences that have been brought to paper and can be shared by sending the according documents (Jensen et al., 2007). Contrarily, tacit knowledge is almost impossible to codify as it typically represents knowledge we are not able to articulate (Grimaldi and Torrisi, 2001). Riding a bicycle exemplifies such knowledge. Explaining to someone, in words, how to balance a bicycle seems impossible; people need to try it themselves and learn from their own experiences. Therefore, tacit knowledge tends to be more spatially sticky (Morgan, 2004), concentrating in particular regions and resisting diffusion (Howells, 2002).

Whether a technology is stronger based on codified or tacit knowledge can be influenced by their structural composition, based on the subcomponents of a technology. An internal combustion engine car consists—among other things—of an engine block, power train and tires. The engine block itself is a technology with several subcomponents, e.g., the cylinders or oil galleries. These components are directly or indirectly connected with each other and their interplay creates a technological system (Arthur, 2009). These systems can be simply or complexly structured. Complex technologies are characterized by large numbers of components that are connected in ways that requires a great deal of information to communicate and understand the structure (Simon, 1962). Moreover, with higher levels of complexity, the likelihood of technological knowledge being codified decreases (Broekel, 2019). Therefore, the difficulties and efforts to understand and advance technologies rise with the levels of complexity (Cohen and Levinthal, 1990; Singh, 1997).

Consequently, not only the generation of technologies but also its diffusion is shaped by complexity (Sorenson et al., 2006). Complex technologies are more difficult to explain and, hence, tend to resist diffusion and concentrate in only a few places (Camagni, 1985; Feldman, 1993; Glückler, 2007). Although the notion of complexity has already been touched upon in theoretical contributions about diffusion, empirical evidence about the relationship between complexity and diffusion is scarce. Exceptions are Sorenson et al. (2006) and Feldman et al. (2015), as well as Balland and Rigby (2017), who analyze the diffusion of complex technologies by incorporating the dimensions of proximity (Boschma, 2005). Balland and Rigby (2017) find geographic proximity to be relevant for the diffusion of technologies and even more so for complex technologies. Contrarily, Feldman et al. (2015) observe a slightly different picture. They divide the diffusion into two phases and find geographic proximity to be irrelevant in the initial phase. Only in the second phase can a distance-driven diffusion be observed. These conflicting results expose that we lack a clear understanding of the diffusion mechanisms of (complex) technologies. Accordingly, this thesis shall complement the existing literature.

The first empirical chapter of this thesis will tackle this gap and offers a novel contribution to the field by incorporating the recent literature on complexity and proximities with the work of Hägerstrand. We will explore in detail which kind of spatial diffusion patterns (Hägerstrand, 1952; 1965) will be adopted by complex technologies. Do complex technologies diffuse hierarchically from city to city and then to neighboring regions? Or do they diffuse contagiously, like a wave from the innovator region? Besides these two diffusion patterns, Chapter 2 will discuss and explore leap-like diffusions. In contrast to hierarchical diffusions, technologies may jump from one region to the other without any tendency of neighborhood effects. Thereby, the concept of Hägerstrand is extended by a third pattern. Differentiating between these patterns will allow us to clarify the role of geographic proximity in the diffusion of complex technologies. Additionally, we will incorporate technological and social proximity to obtain a more complete picture of the diffusion processes. By this, we will add further empirical research to the notion of complexity in EEG and answer the first research question:

Research Question 1: How do complex technologies diffuse in space?

1.2.2 Networks

Based on the seminal paper of Jaffe et al. (1993) which introduces the analysis of patent citations as a “paper trail” of knowledge diffusion, numerous researchers explored and presented the influence of proximities on spatial knowledge diffusion (Peri, 2005; Maggioni et al., 2007; Hoekman et al., 2009; Paci and Usai, 2009). For example, Almeida and Kogut (1997) analyze the moderating role of firm size on geographic proximity and find large firms to cite more distant knowledge sources. Evidence for technological and organizational proximity to facilitate knowledge diffusion was provided by Jaffe and Trajtenberg (1999), who find that patents are more likely to cite each other if they are assigned to the same firm and patent class.

Besides the notion of proximities, existing empirical evidence suggests that social networks facilitate the generation of knowledge by enabling its diffusion (Young, 2000; Montanaria and Saberi, 2010; Schlaile et al., 2018; Tsouri, 2019). The generation of knowledge requires economic actors to utilize internal and external knowledge sources (Lundvall and Johnson, 1994). Only by combining these two sources can a sustained generation of knowledge be achieved; otherwise, economic actors tend to be caught in technological lock-ins (Grabher, 1993). To have access to external knowledge, this must diffuse across space or, in other words, has to be shared between actors (Witt et al., 2012).

This exchange tends to occur while regional actors cooperate in R&D alliances, joint research projects or co-inventorships. During the time of cooperation, technological knowledge is intentionally or unintentionally shared between partners, initiating the possibility to combine this new knowledge with existing technologies and thereby develop innovations. Research in EEG shows that such partnerships tend to be formed between economic agents that share related knowledge, are embedded in similar social, organizational and institutional backgrounds and are located in the same place (Balland et al. 2013; Breschi and Lissoni, 2005; Broekel and Hartog, 2013b; Ter Wal, 2014). Boschma and Frenken (2010) summarize these tendencies as levels of proximity shaping the formation and, thus, the evolution of knowledge networks.

Besides this bilateral or dyadic level, Glückler (2007) emphasizes the role of nodes and the structural level on the evolution of knowledge relations and networks. In this context, nodes represent the entities connected by links, e.g., individuals, organizations or regions. An example of characteristics of the node is the size of organizations, which may determine how many links it is able to maintain at the same time. Larger firms tend to have greater capabilities for actively managing collaborations (Tether, 2002).

In addition to node characteristics, the network structure matters for the evaluation of knowledge relations and networks. In this context, structure refers to the overall composition of networks. In this regards, a typical reference is the “small worldness” of networks (Watts and Strogatz, 1998). The structure of these networks is characterized by many clusters that possess numerous connections between their nodes and a couple of “bridges” to other clusters. Inter-regional knowledge networks can form in such a way when organizations are closely cooperating within their region and at the same time foster good relationships with organizations outside their home region (Schilling and Phelps, 2007). These networks then tend to foster the spatial diffusion of knowledge and technologies.

In accordance with these findings, European policy makers try to foster inter-organizational cooperation through the subsidization of joint research projects. Since 1984, the EU has established eight Framework Programmes with a total volume of 250 billion euros, aiming to fund joint R&D projects between organizations in different EU member states. Additionally, nations like Germany have their own funding schemes with which they support joint projects. For example, in 2008, the Federal Ministry of Education and Research spent approximately nine billion euros on these instruments (Broekel and Graf, 2012). Due to this popularity and financial commitment, it is crucial to critically review these instruments (Bode, 2004; Fornahl et al., 2011; Broekel, 2015).

Accordingly, a vast body of literature has been published, analyzing the effectiveness of subsidies. They mostly find positive effects of subsidized R&D projects on firms’ innovation performance; recent examples are Schwartz et al. (2012), Di Cagno et al. (2016) as well as Czarnitzki and Hussinger (2018). Typically, the innovation performance is assessed by the number of patents generated after the project. It is argued that inter-organizational R&D cooperation leads to knowledge diffusion between partners, subsequently enhancing the number of granted patents. Therefore, patent output is an attractive indicator to measure the effect of subsidized joint projects. It also has, however, one major limitation: it remains unclear whether knowledge has actually diffused between partners. The greater innovation output of firms might have different reasons, for example, the new patents might stem from hiring additional R&D personnel independent from the subsidized project. Thus, the enhancing effect of subsidization is subject to interpretation and we lack a clear picture of whether subsidized R&D projects enhance knowledge diffusion.

The approach developed by Jaffe et al. (1993), using patent citations as paper trails of knowledge diffusion, offers a framework suitable for closing this research gap; it will be applied in Chapter 3. It allows for an analysis of whether subsidized joint R&D projects in Germany have a facilitating effect on subsequent inter-regional patent citations. Thereby, this thesis complements the current empirical literature on the effectiveness of public subsidies and answers the second research question:

Research Question 2: Do subsidized joint R&D projects facilitate the diffusion of knowledge between regions participating in the same joint R&D projects?

The third gap identified in the literature and which this thesis addresses also relates to the network dimension of knowledge diffusion. In the past two decades, economic geographers have put much effort into the analysis of social networks. By understanding their structure and evolution, new insights about the diffusion of knowledge through network relationships have been generated (e.g., Murphy, 2003; Boschma and Ter Wal, 2007; Broekel and Boschma, 2011). Much attention has been directed to the different dimensions of proximity (Boschma and Frenken, 2010) and their effect on link formation. For example, Scherngell and Barber (2009) study the effect of geographical and technological proximity on the formation of collaborative R&D projects. They find both proximity dimensions to be positively related to the occurrence of R&D cooperation. By considering the evolution of industries, Ter Wal (2014) analyzes how the effect of geographic proximity on link formation changes over time and finds it to decrease.

Despite this comprehensive literature, most of these studies exclusively analyze the formation of network links. Although network "[…] variation should be conceived as the results of endogenous mechanisms of network formation and dissolution" (Glückler, 2007, p. 627), the latter part, dissolution, has been vastly neglected in the empirical literature of economic geography. A major reason might be the requirement of longitudinal data, including the formation as well as dissolution date, that is seldomly available (McPherson et al., 2001). Outside geography, some scholars have explored dissolution processes. For example, Polidoro et al. (2011) analyze different forms of embeddedness and their effect on the stability of inter-organizational alliances. Makino et al. (2007) study the factors influencing the intended and unintended dissolution of joint ventures. Still, the research on dissolution processes is scarce.

This contrasts with the importance of link dissolution as a major barrier for knowledge diffusion. Inventors or organizations who stop collaborating or interacting in any way with each other no longer exchange knowledge. Hence, it is vital to build an understanding about the reasons and mechanisms of link dissolution in order to obtain a comprehensive picture of knowledge diffusion. For this reason, this thesis will discuss and analyze factors affecting the dissolution of network links. As proximities strongly facilitate the formation of links at the dyad level, it appears worthwhile to analyze whether their effect holds true for dissolutions as well. For example, economic agents with similar technological backgrounds (i.e., high cognitive proximity) find it easier to communicate and build up a common understanding (Nooteboom et al., 2007). Thus, they are more likely to establish a partnership. However, this cognitive proximity increases the chances that they work in the same industries and may be competitors (Boschma, 2005). The same cognitive proximity that first led to link formation may then also enhance link dissolution, making the relationships unstable (Boschma and Frenken 2010). Therefore, it is expected that the effects of geographic, cognitive, social, organizational and institutional proximity to differ with regards to link formation and dissolution. Additionally, factors at the node and structural level will also be taken into account to answer the third research question.

Research Question 3: Do proximities influence the dissolution of knowledge network links?

1.2.3 Regional context

Thus far, the technological and network dimensions of knowledge diffusion have been discussed in light of how the complexity of technology may shape spatial diffusion patterns and how networks serve as channels of knowledge flows. This leaves the question of whether and how regional contexts shape the generation and diffusion of knowledge.

The European Regional Competitiveness Index 2019 (RCI) indicates that the generation of knowledge, and hence innovation, is concentrated in few regions of the European Union (Annoni and Dijkstra, 2019). Therefore, regions appear to have different possibilities and capabilities to innovate. Feldman (1993) argues that this picture is the outcome of innovation production being a place-specific process. The generation of new knowledge rests on the knowledge already existing and available to people; hence, it is a cumulative process. Accordingly, innovation production concentrates in space. A region that has once successfully innovated tends to continue, as is it successively creates innovation-related capabilities and experiences that helps to innovate even further; a reinforcing process with increasing returns occurs (Arthur, 1990).

With the rise of modern information and communication technologies, this knowledge may be effortlessly transferred between regions, even those at great distance (Friedman, 2005). As a result, regions would converge, meaning their innovation capabilities and new technologies could flourish anywhere in the world. The map of regional competitiveness would become flat without mountains of economic dominance. However, as the RCI already indicates, the world has not become flat; to the contrary it is very spiky (Rodríguez-Pose and Crescenzi, 2008). The diffusion of knowledge shows crucial distance-decay effects (Jaffe et al., 1993) even today. Just recently, Balland et al. (2020) show that the production of complex patents concentrates in a few US regions, mostly large cities. Consequently, the regional context still shapes the generation and diffusion of innovations.

Triggered by an initial innovation, such as modern wind turbines, new industries may emerge when these technologies are adopted by entrepreneurs and businessman (Theyel, 2012). In this way, emerging industries represent the result of an initial technology generation or diffusion. The likelihood of observing the creation of new industries tends to be shaped by regional characteristics. In this regard, Boschma and Frenken (2011) argue that the likelihood of new industries to emerge in regions increases when related industries are present. This presence positively shapes the regional context as the availability of required skills, human capital and infrastructure rises (Boschma, 2017). Empirically, Montresor and Quatraro (2019) find evidence for this argumentation in the case of green technologies in Europe. Subsequently, these regional characteristics are also shaped by the new industry, influencing whether and which innovations are adopted in the future. This interdependence motivates us to take a deeper look at the mechanics of industry emergence in order to understand the influence of regional characteristics on knowledge diffusion.

Many concepts used in EEG, like agglomeration externalities, path dependence or windows of local opportunity, that explain industry emergence focus on the role of the supply-side shaping regional contexts, e.g., access to resources and competencies. The demand-side has received much less attention. However, economic markets are characterized by the supply of products through manufacturers and the demand for these products by consumers (Brem and Voigt, 2009). Firms will only be successful if they match their supply (e.g., what they produce and how much of it) to the actual demand. Therefore, it is necessary that they acquire knowledge about market trends and consumer preferences (Martin et al., 2019). Consequently, close user-producer interactions are necessary for suppliers to correctly recognize future demand and translate it into innovations (Lundvall, 2008; Menzel and Fornahl, 2009). Thus, firms might tend to locate themselves in the vicinity of consumers to identify their demand correctly. Demand is also determined by the information and knowledge consumers have, for example, about new trends in fashion or product features advertised by firms. Hence, actual demand varies locally (Porter, 1990; Justman, 1994) and the magnitude as well as the kind of demand is likely to impact the emergence new industries (Martin et al., 2019). Nevertheless, economic geographers have widely neglected this important driver of change and have focused most of their attention on supply-side factors (Boschma and Frenken, 2011). This motivates this thesis to contribute a missing discussion and elaboration on the effects of demand on the emergence of industries.

Research Question 4: Does local demand shape the spatial emergence of industries?

1.3 Overview of the chapters

To present answers to these four research questions, this thesis will apply an evolutionary framework to all stated dimensions of knowledge diffusion: technologies, networks and regional context. The following sections give an overview of the thesis and introduce the individual chapters by presenting motivation, empirical setup and main results. In addition to the four empirical chapters, the thesis will conclude with Chapter 6, discussing the main empirical results, theoretical as well as methodological contributions, prospects for future research and policy implications.

1.3.1 Technology complexity and spatial technology diffusion

Controlling technological complexity is perceived as economically beneficial, as it promises high returns and offers the possibility to create a competitive advantage due to the difficulties of imitating and copying complex technologies (Balland and Rigby, 2017). For example, Sbardella et al. (2018) discover that economic complexity facilitates economic growth. Besides these benefits, complexity also implies greater constraints and higher resource requirements in its addoption. If a technology has several requirements that need to be fulfilled for adoption, only a few individuals or regions may be capable of its application. Accordingly, it is expected that technologies diffuse with different spatial patterns—hierarchical, contagious (Hägerstrand 1967) or leap-like—when complexity is considered as a technological characteristic. In this regard, Chapter 2 contributes to our understanding of spatial diffusion patterns and how the complexity of technologies shapes the probability of observing either hierarchical, contagious or leap-like diffusions.

In order to understand how technological complexity shapes spatial diffusion patterns, a Bayesian survival framework (Zhou and Hanson, 2017) is applied to a novel complexity index (Broekel, 2019) and geocoded data of 4,000,000 US patents, granted from 1836 to 2010.

Through the calculation of 285 Bayesian survival models, considering geographic, technological, and social proximity as well as population size and technological diversity, the spatial and temporal diffusion of 285 technologies is explored in Chapter 2 and accompanied over a time period of a hundred years. Afterwards, a meta-regression analysis provides insights to how technological complexity affects the strength and relationship of each variable.

The Bayesian survival models reveal spatial diffusion patterns similar to Hägerstrand (1967). Some technologies diffuse contagiously from the region of creation and other technologies spread hierarchically, jumping to distanced regions and diffusing contagiously from these. Still others leap from region to region without showing any neighborhood effect. Additionally, the meta-regression reveals that with higher levels of technological complexity, the likelihood of observing contagious diffusions significantly increases. Further evidence is provided that technological relatedness and diversification facilitate technology adoption, even more for complex technologies. Population size generally facilitates the adoption of technologies, however, less so for complex technologies.

With regards to social proximity, the results show a general tendency to support the diffusion of technologies, but the picture is very diverse. For some technologies the factor has a positive relationship, for others a negative relationship to diffusion speed. These contradicting results bring us to the second dimension of knowledge diffusion, networks, and supports the course of this thesis to a stronger elaboration of the effects and evolution of knowledge networks.

1.3.2 The effect of subsidized R&D networks on knowledge diffusion

From a national or regional point of view, knowledge diffusion between organizations is a favorable process, as it allows organizations to obtain knowledge they would not be able to generate on their own (Powell et al., 1996). This tends to increase the potential for innovations which may subsequently lead to new growth and employment. From the viewpoint of individual organizations, however, knowledge exchange is a double-edged sword; as knowledge leads to competitive advantages, organizations tend to restrain knowledge from partners or even try to exploit partners through opportunistic behavior (Williamson, 1973; Gulati, 1998). In this situation of uncertainty about partner behavior, knowledge might not diffuse without restrictions, and the positive effects associated with knowledge spillovers will probably not occur. Therefore, it appears legitimate that national and regional governments try to foster mutual knowledge diffusion between organizations. For example, in Germany, the government subsidizes joint projects in which organizations are urged to share project-relevant knowledge between all partners (Broekel and Graf, 2012). Otherwise no funding is granted.

Still it is unclear whether these programs lead to the desired, more intense knowledge exchange or to organizations merely using the opportunity to finance their projects partly with public money. Chapter 3 follows the work of Jaffe at al. (1993) and makes use of patent citations as a "paper trail" of knowledge diffusion (Peri, 2005). More precisely, the chapter analyzes whether a positive relationship between policy networks and inter-regional patent citations can be observed. Therefore, patent data from 2000 to 2009 and data from the German "subsidies catalogue" including all supported projects is processed and analyzed in a gravity model framework (Isard, 1954). Besides reflecting on the impact of a popular German policy tool, Chapter 3 also allows the extension of our knowledge on inter-organizational relationships and their impact on knowledge diffusion.

Similar to previous research, Chapter 3 finds a negative relationship of geographic distance to regional patent citations (Jaffe et al., 1993). Moreover, evidence for the facilitating influence of technological proximity is provided. A significant relationship between policy network links and following patent citations could not be found. However, the chapter shows a positive relationship between co-inventor relationships and patent citations. Therefore, the results generally support the diffusion-enhancing effect of knowledge networks. In order to better understand the differences between networks and their effect on knowledge diffusion, Chapter 4 will elaborate on the formation and dissolution of network links.

1.3.3 Proximities and the dissolution of network links

Already in the 1960s, Hägerstrand (1965) concludes that innovations diffuse "through the network of social contacts" and the "analysis of diffusion of innovation may […] be broken down into two parts: the study of links and the study of nodes" (1965, p. 27). However, Hägerstrand treated communication channels, i.e., network links, exogenously (Blaikie, 1978). Thus, he did not further elaborate on the differences of links, network structure or evolution. Because of recent methodological advancements, the mechanisms and factors shaping the structure and evolution of networks can be investigated.

The evolution of networks is characterized by the processes of node appearance and disappearance as well as of link formation and dissolution over time. Studies in economic geography have already thoroughly analyzed the formation of network links. In particular, the work of Boschma (2005) led to several studies analyzing how proximities shape the evolution and structure of networks (Broekel and Boschma, 2011; Balland, 2012; Balland et al., 2015). Besides the relational level, the node and structural network levels have been analyzed as well (Glückler, 2010; Broekel and Hartog, 2013; Ter Wal, 2014). Consequently, a substantial body of empirical evidence of the formation processes has been established, whereas link dissolution was mostly neglected as the required longitudinal data, including dissolution times, is often lacking.

Chapter 4 utilizes the recent methodological advances in the area of exponential random graphs that allow the analysis of formation and dissolution processes of spatial knowledge networks (Krivitsky and Handcock, 2014). Thus, it explores the possibilities of separable temporal exponential random graph models (STERGMs) in order to analyze the network evolution of a German biotechnology network. More precisely, the formation and dissolution processes of policy-induced joint projects from 1998 to 2013 are examined. The data within the subsidies catalogue includes the date of establishment as well as project length. Thereby, Chapter 4 aims to increase the knowledge of the characteristics of inter-organizational relationships in order to gain a better understanding of which relationships are beneficial for the diffusion of knowledge and which are not.

With regards to link formation, evidence is found that geographical, cognitive and institutional proximity increase the chances of link formation between two organizations. In addition, at the node level, support is given that larger firms have greater resources for building and maintaining inter-organizational relationships, as link formation is more likely for large firms. Additionally, the analysis of link dissolution reveals that urban organizations dissolve relationships significantly faster than rural ones, and institutional proximity facilitates link dissolution. Indicating that partnerships between companies of the same organizational backgrounds, e.g., two non-profit organizations, dissolve links faster than relationships of companies with different backgrounds.

1.3.4 Local supply and demand and the emergence of industries

In recent years in EEG several studies have analyzed the origins of new industries. In order to explain why some regions can diversify into new technological systems and industries and others cannot, the evolutionary principle of path dependence is considered (Garud and Karnøe, 2001; Neffke et al., 2011). Organizations invest time, money and human capital into certain technologies and expect financial returns through sold products and services. Over time, economic agents acquire a certain set of skills, infrastructure and routines. According to the technological distance between industries, these skills, infrastructure and routines will differ (Boschma and Frenken, 2011; Essletzbichler, 2015). In their concept of “related variety,” Frenken et al. (2007) extended this argumentation by the dimension of geographical distance and concluded that firms operating in related industries and located in the same regions have the best opportunities to frequently exchange knowledge in a fruitful way. Thereby, they might develop new technologies based on their joint knowledge (Boschma and Frenken, 2012). Accordingly, regions with a broad variety of related industries have the best chances of frequently developing new technologies and, subsequently, industries.

Thus, the recent literature frequently focuses on the supply-side in order to understand the location decisions of industries. Regional capabilities, infrastructure or technological relatedness are used to explain why certain industries emerge in particular regions. The demand-side and its effect, however, have been neglected. Nevertheless, consumer and user preferences are likely to shape the diffusion process as much as the actions of suppliers (Geels, 2002; Martin et al., 2019; Ormrod, 1990). In the end, adopters decide whether they want to use a technology. Chapter 5 will close this gap by simultaneously analyzing supply-side and demand-side factors influencing the diffusion of the German wind energy industry.

Therefore, Chapter 5 takes an intensive look at the interplay of local supply and demand over the life cycle of the German wind industry from 1983 to 2010. On the one hand, it makes use of a Bayesian survival framework to investigate whether and how fast regions deploy their first wind turbines and, on the other hand, if and when they witness the foundation of a wind energy manufacturer. Thereby, demand is modeled in the form of deployed wind turbines. The more wind turbines a region plans to install, the higher its demand. Thus, in the demand-pull models, the location of manufacturers is explained through future wind turbines that are deployed five years after firm foundation. In the supply-push models, the likelihood of wind turbine deployment is explained through the presence of manufacturers.

After extensive robustness checks, the results confirm the importance of related variety, urbanization and industrial agglomeration for the emergence of industries. In addition, evidence is provided that demand is also a crucial factor for understanding the diffusion of an industry. The results show a higher likelihood for firm foundation in regions where numerous wind turbines are planned to be deployed in subsequent years. With regards to wind turbine locations, support is found that natural conditions of regions, such as average wind speed and availability of free space, most determine the adoption of wind turbines.

The spatial diffusion of simple and complex technologies

Abstract

Recent studies show that regions generating and capitalizing complex technologies possess a competitive advantage and, thus, prosper economically. Far less is known about the spatial diffusion of complex technologies. What are the drivers of complex technology diffusion? In particular, do these drivers differ in comparison to simple technologies? To answer these questions, the chapter makes use of four million US patents and analyzes one hundred years of diffusion of 285 technologies. More precisely, for each technology, a spatial Bayesian survival model is fitted, the results of which are evaluated against each technology's degree of complexity. Our findings confirm that simple and complex technologies diffuse with different spatial patterns. More precisely, complex technologies tend to diffuse contagiously. Additionally, we find out that a diverse set of related technologies in a region enhance the adoption of complex technologies to a greater degree than for simple ones. The results also underline the importance of cities in the diffusion process of complex technologies.

This chapter is single authored by the PhD candidate.

2.1 Introduction

Many studies see the generation and diffusion of technologies across space as key drivers of economic growth and, therefore, as important for explaining differences in spatial economic development (Romer, 1990). For instance, regions that are successful in generating and adopting novel technologies have more chances of transforming these into economically valuable products.

In this context, complex technologies have been argued to be particularly valuable, as they are non-ubiquitous and their knowledge is tacit in large parts (Balland and Rigby, 2017). This offers greater chances for monopolistic rents and consequent economic growth (Broekel, 2019). Consequently, complexity is associated with competitive advantages. For example, Sbardella et al. (2018) emphasize a positive relationship between economic complexity and economic growth. This is confirmed by Boltho et al. (2018), who find East Germany's economic convergence to be related to its ability to produce complex goods. In addition, Pugliese et al. (2017) show that more complex economies had an advantage in starting the process of industrialization.

Despite these advantages, the same characteristics that make complex technologies so valuable also suggest greater difficulties and resource requirements in their generation, adoption, and in consequence, their spatial diffusion. As the exploration of complexity is rather new in economic geography, there are few studies investigating the diffusion of complex and simple technologies: One example is the work of Balland and Rigby (2017). They analyze how complexity relates to the diffusion of technologies by studying citation patterns of patents. In this context, they find geographic distance to be a more important obstacle to citations for complex than for simple technologies. Complementarily, Sorenson et al. (2006) pointed out that social proximity is most important for the diffusion of moderately complex technologies.

Although these studies analyze the diffusion of complex technologies, they do not differentiate different spatial diffusion patterns and whether complex technologies show a distinct form of diffusion. Therefore, based on Hägerstrand (1967) and Hagget (2001) we consider three patterns of spatial diffusion: hierarchical, contagious and leap-like. By combining spatial patterns of diffusion with the dimension of complexity, this chapter extends the previous work and generates a more detailed picture of the diffusion of technologies. Moreover, besides the effect of geographic distance in form of different spatial diffusion patterns, this chapter will also analyze two further forms of proximity: technological and social. This will help us to understand why technologies diffuse at different speeds (Pezzoni et al., 2019), and how they form economic landscapes. Consequently, we seek to answer the following research questions: What factors influence the spatial diffusion of complex knowledge? How does their relevance differ from that shaping the diffusion of simple knowledge? In addition, we apply a long-term perspective, including technologies created in the 19th and 20th centuries. This allows us to follow the diffusion of these technologies from their emergence to broad adoption.

To answer these questions, we make use of the novel complexity index of *structural diversity*, which has been introduced by Broekel (2019). Moreover, we rely on the HistPat data base, which includes about four million individual patents granted by the US patent office from 1838 to 2010. On this basis, in the first step we model the spatial diffusion of 285 distinct technologies using Bayesian survival models (Haiming Zhou and Hanson, 2017). In the second, the relation between the relevance of factors driving technologies' spatial diffusion and technological complexity are assessed using a meta-regression approach.

The results confirm the existence of multiple spatial diffusion patterns ranging from contagious to hierarchical to leaping. Hence, geographic proximity varies in its impact on diffusion between technologies. We also show that some of these differences in diffusion patterns are linked to technologies' degree of complexity. For instance, contagious diffusion becomes more likely with rising levels of complexity. The relevance of complexity is also visible in the relevance of technological relatedness. In this case, we identify that the existence of related competences tends to enhance the diffusion of technologies in general and of complex ones in particular.

The paper is structured as follows. Section 2.2 discusses the nature and structure of technologies as well as their diffusion considering the concept of proximities. In Section 2.3, we introduce the empirical setting, including the employed data and methods. Particular attention will be paid to the employed measure of technological complexity. Section 2.4 presents and discusses the results of the empirical exercise. Section 2.5 concludes the paper.

2.2 The structure and diffusion of technology

The diffusion of new technologies is mostly described as an s-curved process (De Tarde, 1903): In the "initial phase," a limited number of actors try out innovations, and these people are called "innovators" and "early adopters" (Rogers, 2003). If they are satisfied with their experiences, they will spread information about those innovations. A re-enforcing process of adoption and communication may subsequently start, leading to an exponential diffusion of the novelty within a population. After some time, products typically reach the so-called "growth phase." This phase is characterized by an increasing growth rate, which will last until the market is saturated, i.e., until fewer potential adopters exist than actual adopters (Rogers, 2003). Eventually, products reach the "maturity phase" and the adoption curve declines.

Although the s-curved process appears very robust regarding the diffusion of technologies, there are significant differences in the time it takes a technology to leave the initial phase and reach the growth phase. Pezzoni et al. (2018) and (2019) analyze the diffusion of 10,000 technologies from 1985 to 2000. Novel technologies are defined as IPC combinations that have not been connected before. To trace the diffusion of technologies, the number of subsequent patents that use the same IPC combination is counted and mapped to its cumulated distribution. Crucially, these authors find technologies to significantly differ regarding their diffusion slopes. Some reach the growth phase very fast, while others require many years or never reach the growth phase.

Pezzoni et al. (2019) focus on the "familiarity" of combined technological sup-components. If two components are assigned to the same upper-level technology (e.g., IPC3) they are more familiar and legitimation time is shorter. Inventors familiar with the upper-level class tend to adopt the new combination more quickly, as they perceive fewer uncertainties. Consequently, those technologies reach the growth phase more quickly. However, they also tend to remain there for less time and enter the maturity phase earlier. New combinations based on familiar components may offer less technological impact as the number of technological applications is smaller (Ibid.).

The present chapter follows the recent literature that suggests complexity is a crucial characteristic of technologies that, among other effects, impacts their diffusion. More precisely, supposing that technologies are systems of directly and indirectly connected subcomponents (Arthur, 2009) allows us to reflect upon their simplicity or complexity, respectively. More complex systems have higher interdependencies between subcomponents, leading to greater difficulties in using such systems as they inherent larger knowledge diversity (Simon, 1962; Kaufman, 1993; Broekel, 2019). Changing one subcomponent might lead to direct or indirect changes in other subcomponents. Thus, besides knowing and understanding all subcomponents, it is also necessary to consider their interdependencies. Only then are successful replication and processing possible. Building up such knowledge can be time and resource intensive (W. M. Cohen and Levinthal, 1990). Therefore, the structure is likely to affect the adoption speed of technologies.

Additionally, as regions tend to develop location-specific sets of skills, technological competencies and institutions (Boschma and Frenken, 2011), it appears likely that regions face distinct efforts and challenges when they try to adopt (complex) technologies. This may then reflect the different diffusion curves identified by Pezzoni et al. (2018) and (2019) and may lead to distinct spatial diffusion patterns. This argumentation follows Balland and Rigby (2017), who argue and confirm the effect size of geographic proximity to vary with the level of technological complexity. More complex technologies are more strongly shaped in their diffusion by geographic proximity, indicating a contagious diffusion pattern. In contrast, Feldman et al. (2015) find that the technology "rDNA"[1] initially jumps from the region of innovation to far distanced cities. Only afterwards can a distance driven diffusion be observed. rDNA is defined as breakthrough technology, i.e., a technology that differs substantially from existing technologies (Phene et al., 2006) and therefore may indicate greater levels of complexity (Rogers, 2003). These contradicting findings about the diffusion of complex technologies reveal that we miss a clear understanding about the diffusion mechanisms of (complex) technologies.

2.2.1 The ambivalent relationship of complexity and geographic proximity

To better understand these patterns, this chapter combines these ideas with the classical literature on the spatial diffusion of innovation. In particular, Thorsten Hägerstrand (1952) describes two effects that shape the spatial aspects of diffusion processes: the "hierarchy effect" and the "neighborhood effect." By empirically analyzing the diffusion of motor cars in Sweden, he observed car adoption to follow a three-stage process. In the first stage, the relative increase in usage is strongest in the innovator region and decreases with distance. In the second stage, other cities adopt the innovation and the relative increase in usage rises with distance. In the third stage, the diffusion speeds converge and the relative increase in usage is similar in all regions. Visually, this process can be understood as a diffusion wave originating from the innovator region (Stage I) and first moving to neighboring regions and then farther away (Stage II). Thereby, the height of the wave symbolizes the usage intensity of the diffusing innovation. After a while, the wave loses momentum and the water, as well as the innovation, has spread evenly (Stage III) (Hägerstrand, 1952).

This implies a contagious diffusion pattern, where nearby agents or regions would always be the first to adopt. However, Hägerstrand (1967) himself argues that the diffusion process is strongly shaped by networks of social relationships. The networks may be geographically shaped, leading to stronger bonds between nearby actors (Howells, 2002), which support the notion of contagious patterns. However, strong relationships might also develop between actors

[1] The recombinant DNA (rDNA) technology based on the Cohen and Boyer patent from 1980, describing the "Process for Producing Biologically Functional Chimeras" (patent number 4237224) (Feldman et al. 2015).

that are proximate in dimensions other than the geographical (Boschma, 2005). In this context, Hägerstrand (1967) argues in favor of cities at the same level of hierarchy having such bonds, leading to a hierarchy effect in innovation diffusion. Expressed differently, from a spatial perspective, cities act as early adopters, “adopting” innovations before their rural counterparts (Brown and Cox, 1971). He defines the hierarchy of cities according to their formal, administrative importance in the political system. Thus, at the top of this hierarchy are global cities and national capitols. Due to their importance in the political system, many firms and institutions have their global or regional headquarters in these cities, leading to strong links between them that are used for extensive information sharing and the early adoption of new ideas. Accordingly, innovations diffuse from the innovator regions to cities at the top of this hierarchy and then down the ladder. From these cities, Hägerstrand (1967) describes innovations spreading to neighboring regions (Fig. 2.1). In other words, not just one wave, emerging from the innovator region will be observable, but several waves that almost simultaneously originate from different cities.

Following Hägerstrand, the role of geographic distance (or proximity) in the diffusion of knowledge or technologies has been the focus of numerous studies in economic geography. A short physical distance between economic actors tends to bring them together and facilitates the exchange of non-codified, tacit knowledge (Boschma, 2005). In other words, close proximity tends to create positive knowledge externalities, i.e., firms benefit from R&D activities conducted by others in the form of knowledge spillovers (Jaffe et al., 1993; Audretsch and Feldman, 1996). This view has been challenged, arguing that geographic proximity is just one of many dimensions of proximities. Boschma (2005) proposes four additional dimensions that affect knowledge exchange: technological, institutional, social and organizational. Before discussing these dimensions, we complete our argumentation about geographic proximities’ effect on the diffusion of complex technologies.

Figure 2.1: Possible spatial diffusion patterns

Simple technologies are characterized by a few distinct combinatorial patterns of only a few technological subcomponents (Broekel, 2019), leading to ease in learning, transmitting, and codifying the underlying knowledge (Feldman, 1993). For example, one can imagine the technological system of a simple technology being structured like a star with one central component which is linked to all other components that have no further relations with each other. In this case, little information is necessary to describe and understand the structure. Consequently, if ignorance about the technology is overcome, adoption should be relatively easy, implying that the diffusion of simple technologies can be expected to be frictionless with geographic proximity being of low relevance. In other words, the technology might leap from one region to another regardless of whether these are closely located or greatly distanced (Fig. 2.1).

Contrastingly, complex technologies present a more diverse structural topology, in which components are stronger and more distinctively interrelated. Such a case is biotechnology, which represents a combination of chemical and biological technologies with additional elements of electronics and information. To adopt and advance this technology, access to laboratories as well as sufficient IT infrastructure, e.g., server farms, is necessary. Thus, the diffusion should be shaped more by the availability of required infrastructure and less by geographic proximity. The likelihood of finding innovation-supporting infrastructure tends to increase with the population size of cities (Bettencourt et al., 2007) and the most populous cities are at the top of the city hierarchy (Hagget, 2001). This suggests a hierarchy-driven diffusion in the case of complex technologies.

On the other hand, as complex technologies tend to consist of more and stronger interrelated components, more information is necessary to describe and understand these technologies (Broekel, 2019). This implies that the sharing of this information is more challenging and less standardized (Feldman, 1993). Actors rely on frequent face-to-face interaction in order to transmit the information, ask questions and discuss the mechanism of complex technologies. These face-to-face interactions tend to occur more often with geographic proximity (Boschma, 2005). This would indicate that complex technologies mostly diffuse contagiously along the social networks of personal interactions (Fig. 2.1).

As we do not know which effect is greater, infrastructure requirements or tacitness, both argumentations might be correct, and therefore two hypotheses shall be pursued.

Hypothesis 1A: With increasing complexity of technologies, the spatial diffusion pattern is characterized by a hierarchical diffusion (i.e., small impact of geographic proximity).

Hypothesis 1B: With increasing complexity of technologies, the spatial diffusion pattern is characterized by a contagious diffusion (i.e., high impact of geographic proximity).

2.2.2 Technological proximity and the diffusion of complex technologies

The adoption of novel technologies in the form of new patents implies that the adopted technology has been advanced (Dosi, 1991), for example by combining it with other (related) knowledge components. In this case, the mere information about a technology seems insufficient, but the ability of recognizing and absorbing it is necessary (W. M. Cohen and Levinthal, 1990). For example, if an actor is experienced in engineering and seeks to master new engineering-based technologies, he will succeed faster than a chemist. Accordingly, experiences with some of the new knowledge's subcomponents will therefore greatly help in adopting technologies. These arguments are at the heart of the technological distance and relatedness debate (Teece, 1981). Thus, though complex technologies may diffuse hierarchically or contagiously, technological proximity is also likely to shape their spatial diffusion.

More precisely, we refer to the idea that technologies may be similar in terms of required inputs, infrastructure, and institutions. That is, development and adaptation is facilitated by the existence of supportive (environmental) conditions for the application of a technology. A region already active in related technologies is likely to have built up a compatible infrastructure and offer required inputs (e.g., human capital), which supports the adoption of related technologies (Boschma and Frenken, 2011). The literature provides numerous studies showing that the emergence and adaptation of technologies by regions is a path-dependent process (e.g., Boschma, 2017; Neffke et al., 2011). Accordingly, regions are more likely to and successful in diversifying into new technologies when these are related to pre-existing activities (Boschma, 2017). Balland and Rigby (2017) empirically support this argument by finding technology relatedness to enhance the probability of patent citations. We argue in this paper that the role of technological proximity is even more crucial in the context of the diffusion of complex technologies. The greater heterogeneity of knowledge will increase the benefits of compatible infrastructures and, as shown above, actors' abilities to learn and utilize it. Recall the biotechnology example; technological proximity to either one or even both components—chemistry/biology and electronics/information—appears beneficial when adopting this complex technology.

Hypothesis 2: A diverse set of related technologies in a region is more relevant for the diffusion of complex than of simple technologies.

2.2.3 Complexity diffusion along social relationships

Two steps are necessary for successful technology transfer: first, becoming aware of the technology and, second, understanding its structure and mechanisms. For the first, any kind of information channel may be sufficient, from mass media to personal communication (Ormrod, 1990). In the case of the second, understanding technologies, personal interactions seem to be more appropriate. This is particularly true for complex technologies with their components belonging to different knowledge bases, which hinders learning and understanding (Dodgson, 1992). In addition, complex technologies tend to embody greater parts of tacit knowledge (Sorenson et al., 2006; Balland and Rigby, 2017).

Developing technologies and designing products on this basis, for example in biotechnology, requires the expertise of a diverse set of actors (Pavitt, 1998). By combining their knowledge, they jointly develop the capability to successfully merge these heterogenous components. Such an exploitation and combination of distinct knowledge bases is a sophisticated and uncertain process that tends to exceed the capacities of individual organizations that are specialized in specific technologies and activities (Kirkland, 1961; Nelson and Winter, 1982). Consequently, Powell et al. (1996) find that in biotechnology the “locus of innovation” is located in a network of inter-organizational relationships, i.e., the combination of the distinct knowledge bases is achieved in inter-organizational learning processes frequently organized in collaboration.

Economic agents are embedded in systems of social relationships, e.g., kinship or friendship (Granovetter, 1985). The strength of this embeddedness can be described by the notion of social proximity (Boschma, 2005). The more economic agents communicate with each other, for example through joint technology projects, the more proximate they are to each other. This embeddedness tends to fuel the development of trust between actors because, based on previous experiences, they can judge each other better and they may share similar values (Ibid.). Thus, the heterogeneity and tacitness of knowledge in complex technologies is likely addressed by inventors collaborating with socially proximate partners.

Hypothesis 3: Social proximity is more relevant for the diffusion of complex than of simple technologies.

2.3 Empirical setting

We are interested in if and how the diffusions of simple and complex technologies differ. For this, we adopt a two-stage procedure. In the first stage, we calculate Bayesian survival models that evaluate the time required for a region to adopt[2] a new technology. In this analysis, we consider the processes discussed above and further factors that influence the speed of technology diffusion. In a second stage, we perform a meta-analysis on the first-stage results. More precisely, we calculate a linear regression model with the technologies' degrees of complexity as the meta-independent variable and the focal first stage variables' coefficients as meta-dependent variables (Jarrell and Stanley, 1989). On this basis, we gain insights into whether the relevance of regional factors in the diffusion of technologies systematically varies in relation to technologies' degrees of complexity.

2.3.1 Bayesian survival analysis

Here, the adoption and diffusion process is modelled between regions and starts with a second region becoming active in a specific technology. The adoption of new technologies by regions is an event which takes place at a given moment in time. The likelihood of observing this event is shaped by the capabilities of regions making sense of new technologies as well as their embeddedness in the inter-regional innovation systems. To assess which factors influence the time required from the "outbreak" of a new technology in one place to its adoption in another, we use Bayesian survival models[3].

Survival models were developed in the field of medical research and sought to explain the risk of patients falling ill or even dying. This risk is perceived to be conditioned by several covariates of theoretical interest, e.g., the physical condition of the patient (Fox and Weisberg, 2011). Survival models have been also adapted to economic geography. For example, Feldman et al. (2015) make use of a Cox survival model in order to explain the diffusion of rDNA technology.

Survival models are generally constructed as follows:

$$p_{(t)} = p_0(t) \exp\left(\beta^T x(t)\right)$$

[2] Regarding terminology, the region which is the first to patent a new technology is called the "innovator region" and all regions that patent afterwards are labelled "adopter regions." In this case, adoption refers to the process of becoming aware of a technology, understanding it and developing it further.

[3] Survival models are also known as "event-history analysis" in sociology or "failure-time analysis" in engineering.

Where $p_{(t)}$ is the likelihood of an event at time t (e.g., the approval of a patent), $p_0(t)$ is the exogenous baseline hazard, i.e., the likelihood of observing an event independently of any further covariates. $x(t)$ is a vector of variables (e.g., proximities) which probably affect the baseline hazard, and β^T represents the according covariates (Perkins and Neumayer, 2005). In this paper, the event is defined as the point in time when a region has its first patent granted in the specific technology.

In comparison to standard regression models, survival models consider the effect of censoring in longitudinal data. Events may lie outside of the observation period, i.e., they might have happened earlier (left-censored) or later (right-censored). Standard regression does not take censoring into account and thereby misjudges the time it takes until an event occurs (Mills, 2011). The Bayesian version of these models allows the consideration of random effects or so-called frailties (Darmofal, 2009). If not considered, these frailties may lead to underestimation (overestimation) of the factors positively (negatively) influencing the hazard rate (Box-Steffensmeier and Jones, 2004). In case of georeferenced data, it is assumed that frailties may result from a Gaussian random field (GRF) (for a detailed discussion, please see Zhou and Hanson (2017)). In the present paper, we make use of geolocational data, that is, we model potential spatial dependencies using information on regions' central geo-coordinates and their distance to neighboring ones.

2.3.2 Meta-regression analysis

Individual survival models are estimated for a set of technologies. To explore to what extent variations in the calculated coefficients are related to technologies' degrees of complexity, we feed the survival regression results into a meta-regression. The approach of meta-regression analysis (MRA) originates from the literature seeking to quantitively review and summarize potentially varying results of studies on the same topic (Jarrell and Stanley, 1989). Especially in medical or psychological sciences where repetitive (clinical) studies are conducted, MRA can help to identify the reasons for varying test results. In this context, the studies' results are evaluated with respect to a "meta-independent variable." In the context of medicine, this is usually the country in which the studies have been conducted or the sample sizes of the individual studies.

In this paper, we conduct the MRA with the set of individual diffusion (Bayesian) models as our "units of observation." On the basis of this set, we calculate a linear regression model for each explanatory variable in the diffusion models. The (beta) coefficients obtained in the diffusion models, e.g., that of geographic proximity, serve as observations. To these, we match the corresponding complexity values of the technologies, which serve as the meta-independent variable. Accordingly, we seek to explain variations in the magnitude of the coefficients with the technologies' levels of complexity.

2.3.3 Survival data

To study the diffusion of technologies, we first need information on technologies. We follow existing approaches in the literature and use patent data as well as its classification by the Corporate Patent Classification (CPC). The CPC was jointly developed by the European as well as US patent offices and was introduced in 2013. It is organized into nine research areas such as "Human Necessities" and "Chemistry" on the top level and more than 250,000 sub-classes on the lowest level. We define technologies as four-digit CPC classes and accordingly study their diffusion. The four-digit level represents a compromise of technological disaggregation and practicability, as finer levels of disaggregation imply larger numbers of models to be run. In addition, we follow other studies that made use of this level (e.g., Breschi and Lenzi 2012; Schmoch and Laville 2003), allowing for matching information on technologies' complexity (Broekel, 2019).

Using patent data implies a number of limitations. For instance, they capture only parts of the technological knowledge (Criscuolo and Verspagen, 2008) and the probability to patent is also known to differ between industries (Arundel and Kabla, 1998). However, they are still the best indicator of technology generation and diffusion (e.g., Jaffe et al., 1993; Fleming and Sorenson, 2001). This is especially true in the context of the present paper, as there is no other indicator of inventive activities that covers multiple technologies over an extensive period of time.

Our empirical basis is the HistPat data base (Petralia et al., 2016) including 4.8 million US patents from 1836 to 1975 and the NBER Patent Database, which extends this data to 2010. We aggregate the location information of inventors' residence to the level of US Metropolitan Statistical Area as this level is commonly used in the literature (e.g., Balland and Rigby, 2017; Feldman et al., 2015; Jaffe et al., 1993). Subsequently, we create a data set for each CPC containing the first patent granted in each MSA, which represents the adoption events in the diffusion models. This results in 655 data sets representing the diffusion of individual CPCs since 1836, i.e., over a period of 174 years. However, CPCs differ widely in terms of when they have been first introduced, implying that in many cases we do not observe any diffusion in many early years (see Fig. 2.2).

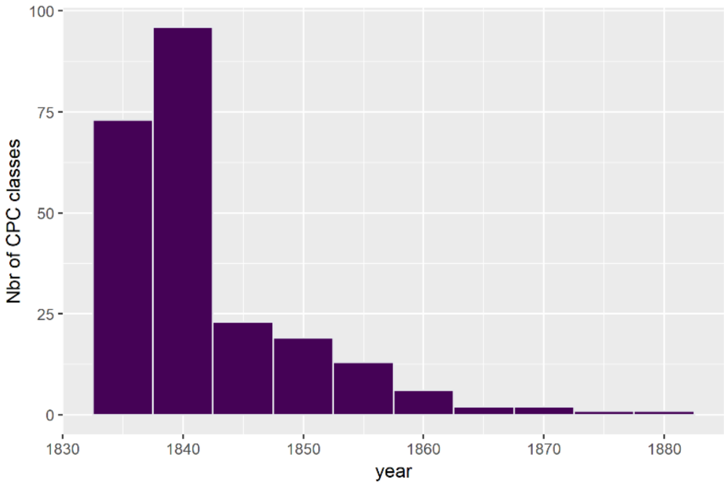

Figure 2.2: Histogram of CPC classes and first year of observation

We address the heterogeneity in observational periods by focusing on the first one hundred years of diffusion for each CPC. In other words, if a CPC was introduced in 1836, we only consider the period from 1836 until 1936. Consequently, we also exclusively consider CPCs that were introduced before 1910. Despite restricting the observational time period, we still cover the majority of four-digit CPCs (471) and their diffusion in the 19th and 20th centuries. While the chosen time frame and restrictions are admittedly arbitrary, they in our eyes offer the best trade-off between comparability, numbers of included CPCs and length of time frame. Due to data availability (regional population), we aggregate the data into periods of ten years.

Finally, we exclude all technologies that have their first occurrence in 1836. This is the first year of our observation period; however, patents have been granted since 1790. Therefore, we cannot identify for how long technologies of 1836 had previously been present. To some more minor extent, this is also true for the subsequent years. In order to minimize this bias, we exclude from our data set all technologies that appear for the first time in 1836.

To increase the chances of stable, converging Bayesian models, we focus on technologies that have been adopted by at least 50 MSAs. This is in line with the literature stating that for each explanatory variable about ten events shall be present (Breul et al., 2015). This implies a significant reduction in considered technologies (from 401 to 285). Obviously, this might induce a bias, as the most complex technologies cannot be expected to diffuse to many regions. We address this by comparing the technologies' complexities of the initial and the sample data set, see Table A.2.1 and Figure A2.1 in the chapter's appendix. The distributions of both sets of technologies' complexity values are very similar. Accordingly, we are confident that we do not introduce any bias with our sample selection.

2.3.3.1 Geographic Distance

To explore the structure and spatial determinants of the diffusion of technologies, we create a number of variables. The first is *DISTANCE.ADOPTER*$_{i,j,t}$, which represents the geographic distance of region *i* to the geographically nearest (Euclidian distance) region *j*, which has already adopted a particular technology in a particular year t. The values of this variable may change over time, as a technology may diffuse to regions that are closer.

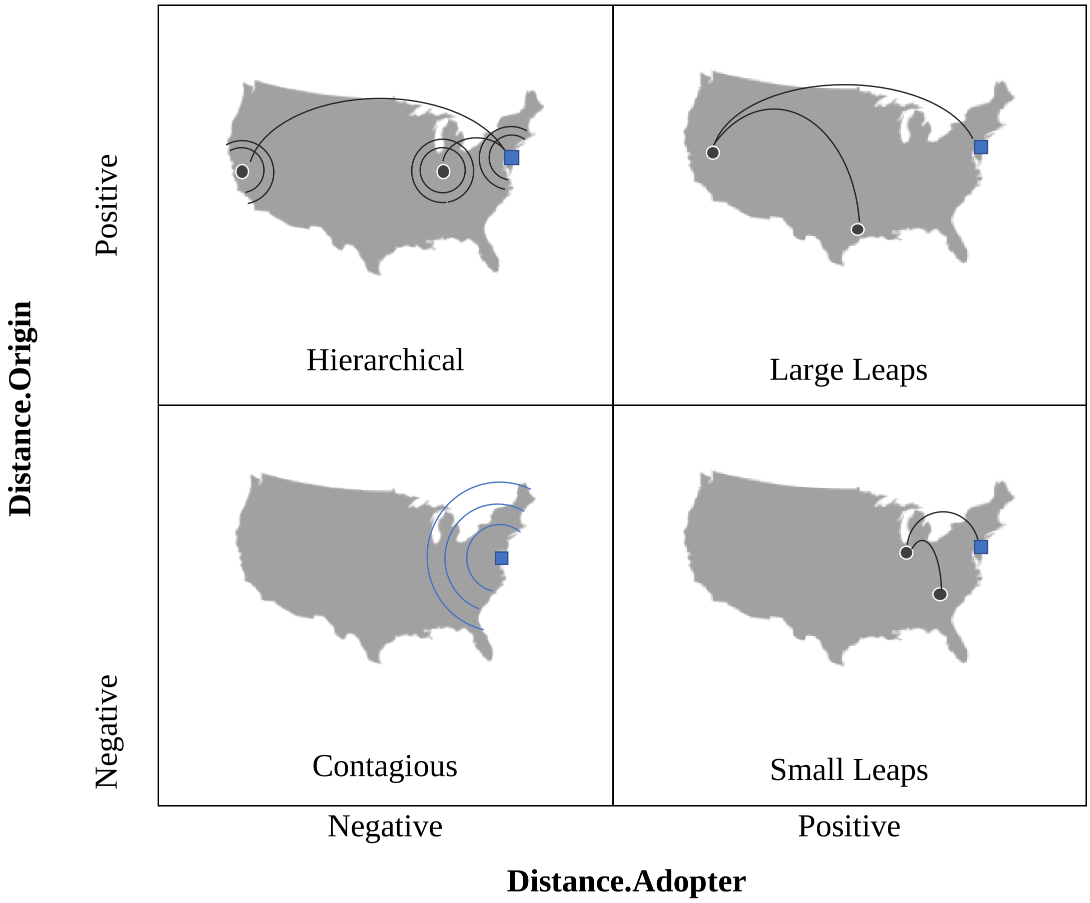

Figure 2.3: Diffusion patterns and geographical proximity

Additionally, we consider the distance of region *i* to the region where the technology was originally invented (*DISTANCE.ORIGIN*$_{i,j}$). In contrast to *DISTANCE.ADOPTER*$_{i,j,t,}$ this variable's values are time-invariant. The combination of both distance variables allows for differentiating between hierarchical, contagious and leap-like diffusion patterns (Fig. 2.3). During a hierarchical diffusion, technologies tend to jump away from the innovator region and, subsequently, diffuse in waves from adopters. Therefore, *DISTANCE.ORIGIN* is expected to be significantly positive and *DISTANCE.ADOPTER* significantly negative. In cases of contagious diffusions, the technology spreads like a wave originating from the innovator region. In this scenario, both distance variables are expected to be significantly negative. When

DISTANCE.ADOPTER is significant positive, no contagious diffusion is observable, indicating that the technology leaps between regions, either a great distance from the innovator regions (*DISTANCE.ORIGIN* becomes positive) or in proximity to the innovator region but in different directions (*DISTANCE.ORIGIN* becomes negative).

2.3.3.2 Social Proximity

We have argued that social proximity may enhance the diffusion of technologies, as socially proximate partners tend to have some shared history which facilitates the generation of trust and, consequently, the reciprocal transmission of knowledge. Therefore, these actors may exchange knowledge about new technologies more frequently, which accelerates diffusion. Such a technology-related shared experience may be the co-inventorship of a former patent. In this case, two or more researchers have worked on the same patent, i.e., technology, and may have built up trust, leading them to keep in touch and share information relating to new technological developments. Hence, to calculate region *i*'s social proximity ($SOCIAL_{i,j}$) to an "infected" region *j*, we evaluate the regions' intensity of joint patent co-inventorships in the prior five years (similar to Ter Wal (2014)). That is, we count the number of patents with inventors from both regions *i* and *j*.

2.3.3.3 Technological proximity

Technological proximity ($TECH_{i,c}$) is the association strength of region *i* and a technology *s*. Here, we derive a 474 x 474 co-occurrence matrix considering the joint occurrence of CPCs in the same patent with the last 5-years. This moving time window is commonly used in the literature (Breschi and Lenzi, 2012; Buchmann and Pyka, 2015). Subsequently, we identify the technology in which region i has already patented that is most similar to the focal technology *c*. The co-occurrence data has been normalized by considering the total co-occurrences. Thus, we prevent the association strength from correlating with the number of occurrences (for a detailed discussion see Van Eck and Waltman (2009)).

2.3.3.4 Regional characteristics

In contrast to the variables introduced above, which capture the relation between regions, we also consider some characteristics of the individual regions. First, we take into account the total population for each MSA and year ($POP_{i,t}$). This allows for control of urbanization effects. The relationship between the diffusion of technologies and city size has already been discussed in the works of Marshall (1920) and Jacobs (1969). Marshall (1920) explained the positive relationship of industrial productivity and city size with better opportunities for input sharing, improved access to work forces and higher probabilities of knowledge spillovers. Empirically, Balland et al. (2020) explored whether complex economic activities concentrate in cities. They find evidence for complex patents, publications and industries to appear more often in highly populated MSAs. Accordingly, we expect more populous regions to be more likely to be early adopters of new technologies in general and complex ones in particular.

Second, regions can be either more specialized or diversified. The former happens when only one or a few industries locate in a region, the latter when several industries are present. Both forms can have positive effects on the adoption of complex technologies. In the case of specialization, it shows that a region is able to develop institutions and infrastructures geared towards a particular industry. In this sense, it might represent a region's "capacity for dedication" towards individual industries or technologies. Being specialized in few industries or technologies bears the danger of lock-ins and strong dependency on one of these, which local actors may wish to reduce and, hence, actively seek to diversify their region (Jacobs 1969). Similarly, highly diversified regions are more likely to adopt new technologies because they have a "proven" capacity of hosting and supporting multiple industries/technologies at the same time. Highly diversified regions signal the capacity to sustain high levels of diversity and, hence, a capacity to provide necessary resources and niches for new technologies to grow. In this regard, Jacobs (1969) argued that diversity might lead to cross-fertilization of technologies and, thus, to more innovation (Jacob's Externalities). Therefore, we construct a simple variable capturing the degree of technological diversity of regions (*DIVERSE*). For each region we count the number of technologies (four-digit CPC) for which a region has been granted a patent in the past five years. We neglect more common and complex diversity measures as we observe only a few patents per region in the 19th century. Accordingly, an index like the location quotient would not be adaptable. Technological diversity of cities is also partly reflected by $POP_{i,t}$ as larger cities tend to have more diverse technology portfolios but can also focus on few technologies (Marshall, 1920).

2.3.4 Structural complexity—the meta-independent variable

A central variable in our empirical approach is technological complexity. To quantify this on the basis of patent data, we rely on the approach of Broekel (2019). Based on the understanding of technologies as systems consisting of several connected components, Broekel (2019) describes technologies as "combinatorial networks" (p. 2) and derives a measure of complexity. The airplane technology, for example, consists of knowledge components such as wing design and aluminum processing that can be thought of as nodes in this network; their interdependence in the airplane forms the network link (Ibid.). Broekel (2019) argues that heterogeneity in these components translates into distinct network structures. As diversity is widely accepted to be closely linked to complexity, he argues that the diversity in these structures (network topologies) is a sign of the overall complexity of the technology. More precisely, if the combinatorial network of a technology is similar to an ordered network (e.g., a star) and hence has a low topological diversity, the technology is rather simple. In contrast, networks of complex technologies are expected to be characterized by higher structural heterogeneity (e.g., small-world networks) and "[accordingly], the more information is required to describe the topology of a technology's combinatorial network, the more complex it is" (Ibid., p. 4). By translating this idea to information contained in patents, Broekel (2019) derives an index of technological complexity (called *structural diversity*) which can be empirically approximated in the context of patent data with the network diversity score of Emmert-Streib and Dehmer (2012).

We follow this approach and calculate the individual network diversity score for each technology's (four-digit CPC class) network ($iNDS_c$):

$$iNDS_C = \frac{\alpha_{module} * \gamma_{graphlet}}{\theta_{module} * \theta_{\lambda}}$$

Here α_{module} is the share of modules, calculated by dividing the number of modules M by the number of nodes n. As we look at undirected binary networks, $\gamma_{graphlet}$ is the share of graphelets of size three and four. θ_{module} measures the variance of module sizes m and finally θ_{λ} is the Laplacian (L) matrix's variability. A network might show its properties by chance; therefore, we calculate the $iNDS_c$ for a set of random sample networks G_M from network c:

$$NDS\left(\{G_c^S | G_M\}\right) = \frac{1}{S} \sum_{G_c \in G_M}^{S} iNDS_C$$

Finally, we calculate the structural diversity of each CPC for each year during the first one hundred years of its existence. To cope with small patent numbers, we use a three-year time window and aggregate the patent data accordingly. In order to assess the complexity of each technology by one value that is usable in the meta-regression, the average complexity for each year over all technologies was calculated and compared with the focal technology's complexity value of that year. If it is above the average, 1 has been assigned, 0 otherwise. Afterwards, the sum has been calculated, ranging from 0–100, according to the years a technology was above average complex (*CPLX*). This procedure allows us to consider the change of complexity over years and creates a single value that serves as the meta-explanatory variable in the second-stage meta-regression approach. Tables 2.1 and 2.2 present the simplest and the most complex technologies, also including the average complexity of a technology over one hundred years ($CPLX_{ARVG}$).

Table 2.1: Top 3 simplest technologies according to the structural complexity index (first one hundred years of diffusion)

CPC-Class	**Description**	**CPLX**	**CPLX$_{ARVG}$**
A63G	Merry-Go-Rounds, Swings, Rocking-Horses etc.	0	2.38
A01L	Shoeing of animals	0	1.46
F41B	Weapons for projecting missiles without use of explosives (e.g. spears)	1	1.79

Table 2.2: Top 3 complex technologies according to the structural complexity index (first one hundred years of diffusion)

CPC-Class	**Description**	**CPLX**	**CPLX$_{ARVG}$**
F03B	Machines or engines for liquids	98	5.99
B29C	Shaping or joining of plastics	97	5.48
E05F	Devices for moving wings into open or closed position	95	6,13

2.4 Results

We visualize the results of the second-stage meta-regressions by means of scatterplots (see Fig. 2.5). In these plots, each dot represents the coefficient obtained for the focal factor in the Bayesian diffusion model for a particular technology. For example, in Figure 2.5A the dots show the corresponding values of each technology class for the variable *DISTANCE.ORIGIN* for each technology class. The color of the dots signals whether the according values are significant or not. The purple line represents the fitted meta-regression, whereby the slope parameter indicates the relationship of the focal factor's importance (size of coefficients) in the diffusion model with the complexity of the underlying technology. A normalization of coefficients is not necessary as the diffusion models calculate the odd ratios for observing an event. In each meta-regression, we fit one variable of the diffusion models as the meta-dependent variable (e.g., *DISTANCE.ORIGIN*) with *CPLX* as the meta-independent variable. This leads to six meta-regressions, which results are summarized in Table 2.3.

In the first step, we interpret the results for the spatial distance variables, which will give us a general overview of the spatial diffusion process. Subsequently, we will focus on the non-spatial variables.

Table 2.3: Descriptive Statistics

Variable	Mean	St. Dev.	Min	Max
β(DISTANCE.ORIGIN)	-0.032	0.039	-0.16	0.08
β(DISTANCE.ADOPTER)	0.01	0.101	-0.328	0.256
β(POP)	0.798	0.203	0.102	1.271
β(TECH)	0.002	0.003	-0.005	0.014
β(SOCIAL)	0.0003	0.001	-0.002	0.002
β(DIVERSE)	0.0001	0.001	-0.003	0.006

The effect of geographic proximity is approximated by two variables, *DISTANCE.ORIGIN* and *DISTANCE.ADOPTER*. The first measures the distance between potential adopter region *i* and the original innovator region. The second represents the distance to the closest region that has already adopted the technology. On average, *DISTANCE.ORIGIN* obtains negative coefficients (see Table 2.3 and Fig. 2.5A). This implies that the relation between diffusion time and distance is negative. This general negative effect of geographic distance on knowledge diffusion is in accordance with numerous other studies of technology and knowledge diffusion (e.g., Jaffe et al., 1993; Bednarz and Broekel, 2019). In addition, *β(DISTANCE.ADOPTER)* has a positive mean (see Table 2.3 and Fig. 2.5B), indicating that technologies rather leap from one adopter to the next instead of diffusing like a wave.

Plotting both distances in a scatterplot gives further details on the spatial diffusion patterns (see Figure 2.4). According to *DISTANCE.ORIGIN* and *DISTANCE.ADOPTER* either being positively or negatively significant, four combinations are possible. Quadrant 1 shows technologies where *β(DISTANCE.ORIGIN)* and *β(DISTANCE.ADOPTER)* both obtain negative values, indicating a contagious diffusion pattern. Quadrant 2 presents technologies with *β(DISTANCE.ORIGIN)* being positive and *β(DISTANCE.ADOPTER)* again negative, implying a hierarchical diffusion. Therefore, technologies in these quadrants diffuse in accordance to the patterns described by Hägerstrand (1967). Quadrant 3 and 4 illustrate technologies with *β(DISTANCE.ADOPTER)* being positive and *β(DISTANCE.ORIGIN)* either negative (Quadrant 3) or also positive (Quadrant 4). The first indicates a leap-like diffusion patterns with a rather short distance to the innovator region and the second one implies large leaps. In line with the before-mentioned mean value of *β(DISTANCE.ORIGIN)*, Figure 2.4

shows most dots in the negative value range, suggesting that a high geographical proximity to the innovator region is beneficial for fast adoption. In addition, by also including *β(DISTANCE.ADOPTER)*, we can differentiate in more detail and observe three diffusion patterns at hand. Quadrant 4 will be neglected from now on, as only four significant observations are in this area. Consequently, a leap-like diffusion pattern with large distances to the innovator region seems to occur rather seldom.

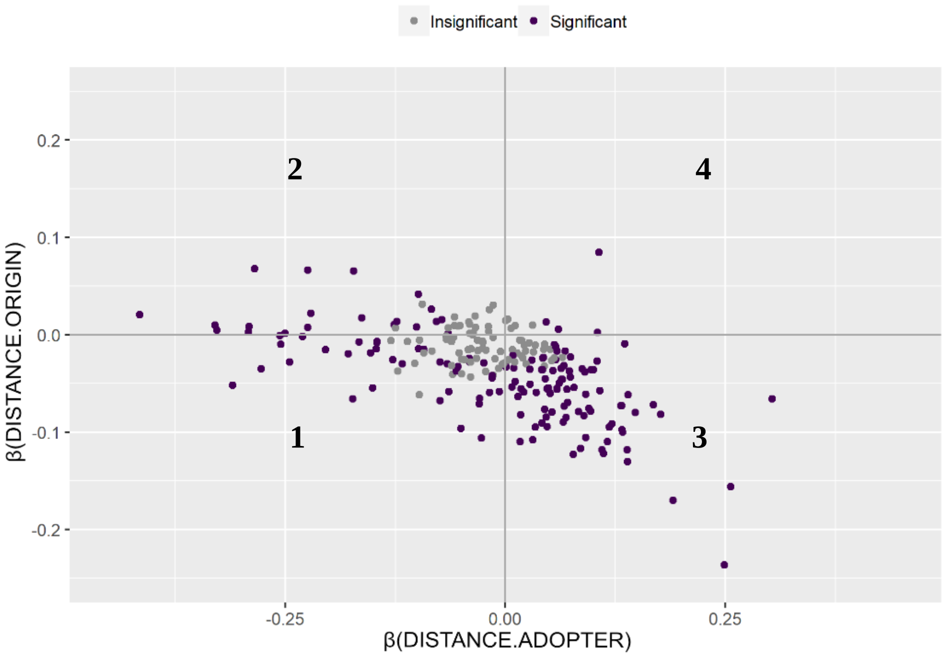

Figure 2.4: Scatterplot of *β(DISTANCE.ORIGIN)* and *β(DISTANCE.ADOPTER)*

Quadrant 3 shows technologies in which both distances are significantly related to diffusion time. However, *β(DISTANCE.ORIGIN)* and *β(DISTANCE.ADOPTER)* obtain opposite signs. The positive sign of *β(DISTANCE.ADOPTER)*suggests a positive effect on the diffusion speed of distance to the closest region that has already adopted a technology. Accordingly, these technologies seem to leap from region to other greatly distant regions. However, distance to the innovator region is also relevant, as *β(DISTANCE.ORIGIN)* is significantly negative. Hence, these technologies accomplish this within the neighborhood of the inventor region. Potentially, this characterizes a simultaneous diffusion "in all directions," i.e., regions that are similarly located close to the inventor region but in different directions. An example of such a diffusion is presented in the appendix, Figure A3.2.

To evaluate whether the occurrence of these patterns is related to technological complexity, three binary dummy variables are created, with either 1, meaning the technology belongs to either quadrant one, two or three, and 0 if not. Afterwards, these dummy variables are considered in three further binomial logit meta-regressions with the dummy variable as the meta-dependent variable and *CPLX* as the meta-independent variable (see Table 2.4). In case of hierarchical and small leaps, the regressions do not provide a significant relationship. Accordingly, technological complexity does not increase or decrease the chances of a technology diffusing either in a hierarchical or leap-like pattern. For cases in which technologies diffuse contagiously, *CPLX* obtains a significantly positive value. Thus, if technologies are more complex, the likelihood of them diffusing in wave-like patterns originating from the innovator regions increases. In other words, this confirms Hypothesis 1B. As complex technologies are composed of tacit knowledge in larger parts, face-to-face interactions are necessary for their transfer. Geographic proximity is likely to facilitate such interactions, especially in the 19th century, as people's mobility was more restricted than today. Cars and planes were not yet invented, and the railroad network had only started to emerge (Stevens, 1926).

Table 2.4: Logic regression of diffusion pattern and technological complexity

	Binomial Regression			
	Coefficient	p-value	N	Positive N
Contagious (Quadrant 1)	0.02***	< 1-04	145	30
Hierarchical (Quadrant 2)	-0.004	0.572	139	9
Small Leaps (Quadrant 3)	-0.01	0.243	150	89

Balland et al. (2020) show that complex technologies generally concentrate in cities and that this urban concentration of complex technologies has increased over the last 150 years. The advantage of cities in initial phases of technologies' diffusion processes is further supported by our results and, more precisely, by *POP,* which tends to be positively related to the time of adoption (see Table 2.3). In other words, the more populated a region, the more quickly it tends to adopt new technologies. Surprisingly, the effect strength decreases with complexity (see Fig 2.5C and Table 2.5). However, the coefficient of *POP* remains positive. Accordingly, while cities have an advantage compared to rural regions when it comes to the adoption of technologies, this advantage is lower for complex technologies. Stated differently, technological complexity appears to somewhat balance the adoption advantage of cities. Some care needs to be taken, however, as the majority of observed technologies began their diffusion before 1850 (see Fig. 2.2). Moreover, concentration tendencies of technologies have been lowest in the mid-19th century and have risen continuously since then (Balland et al., 2020).

Consequently, the advantage of cities in adopting complex technologies was less pronounced at that time.

Table 2.5: Results of the meta-regression models

Model	Complexity	p-value
β(DISTANCE.ORIGIN)	-0.0001	0.358
β(DISTANCE.ADOPTER)	-0.0004	0.12
β(TECH)	0.001***	< 1-04
β(SOCIAL)	0.000002	0.32
β(POP)	-0.004***	< 1-04
β(DIVERSE)	0.0001***	< 1-04

We also control for the level of regional technological diversity by measuring how many technologies a region is active in at the time of adopting the diffusing technology. *DIVERSE* frequently obtains a positive coefficient in the diffusion models (see Table 2.3 and Fig. 2.5D). This indicates that regions patenting in several technologies are more likely to adopt new technologies more quickly. This is in line with Jacobs (1969), who argues that regional diversity might lead to cross-fertilization of technologies and, thus, to more innovations (Jacob's Externalities). This study adds to this that diversified regions not only innovate more but also adopt (complex) technologies more quickly. The meta-regression also confirms a significant relationship between technologies' complexity and the effect size of *DIVERSE* (see Table 2.5)**.** Accordingly, regions with higher levels of diversity are quicker to adopt complex technologies.

This effect is even stronger when regions are active in related technologies as *TECH* mostly obtains positive coefficients in the diffusion models (see Table 2.3 and Fig. 2.5E). A region that is already active in related technologies adopts new technologies earlier than regions that have a specialization in unrelated technologies. This is in line with many previous studies (Feldman et al., 2015; Pezzoni et al., 2018) and supports the working of relatedness (Boschma and Frenken, 2011). With regards to the meta-regression, the effect strength of technological proximity is also positively dependent on technological complexity (see Table 2.5). Thus, technological relatedness becomes even more important in the adoption of highly complex technologies, which is in line with Hypothesis 2. Due to the high structural heterogeneity of complex technologies, regions face greater uncertainties when adopting these technologies. Building upon related capabilities and infrastructure in such cases facilitates the likelihood of a successful adoption. It might even accelerate the decision to adopt a complex technology in the first place. Other regions facing the same uncertainties and having only unrelated expertise will need more time to catch up and successfully adopt complex technologies.

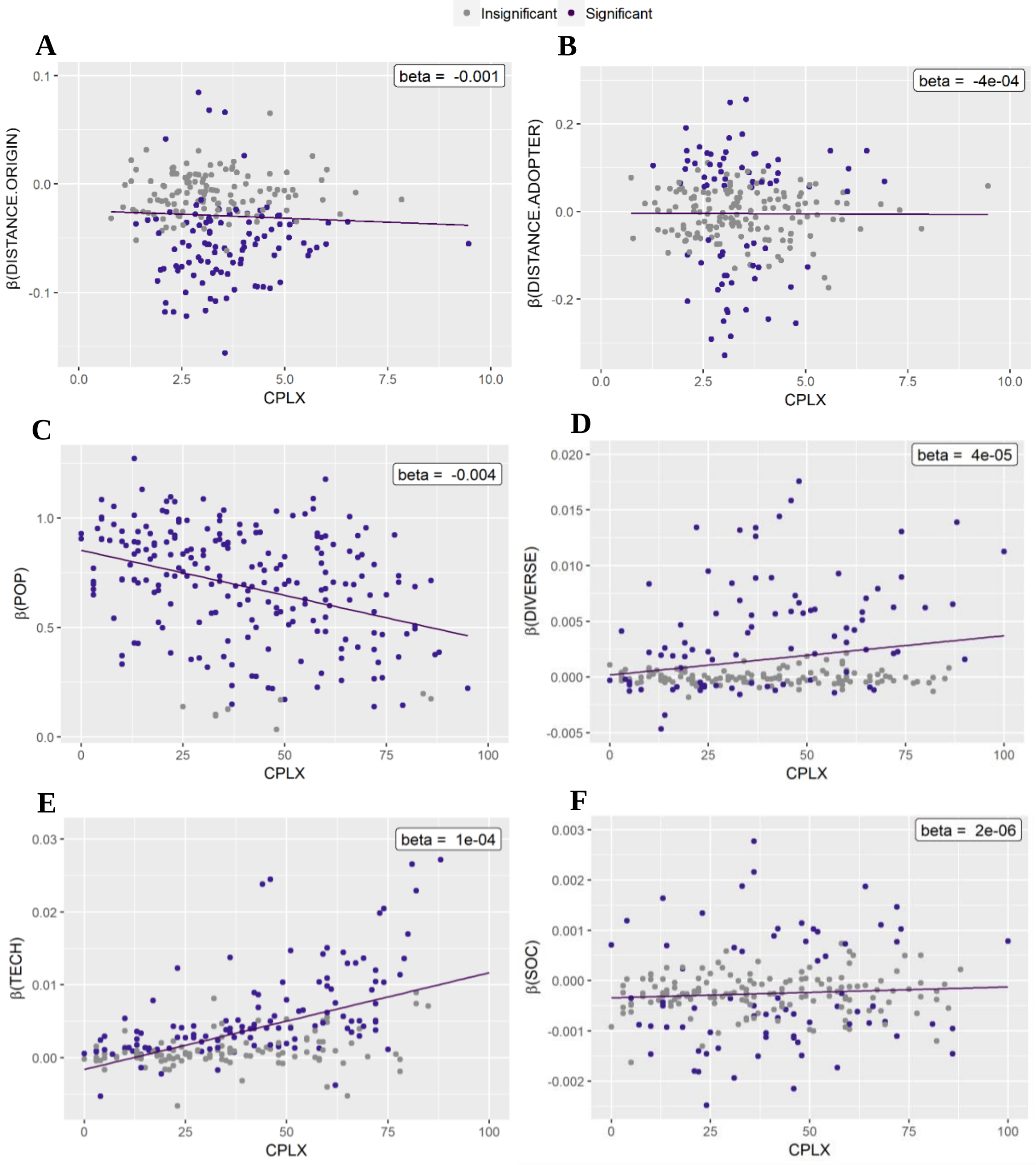

Figure 2.5: Relationship between average technological complexity and **A** β(DISTANCE.ORIGIN), **B** β(DISTANCE.ADOPTER), **C** β(POP), **D** β(DIVERSE), **E** β(TECH) and **F** β(SOCIAL)

With respect to social proximity *SOCIAL* (co-invention), we observe a mix of negative and positive coefficients in the diffusion models (see Fig. 2.5F). The latter meets our expectations. Social embeddedness supports learning and lowers transaction costs, which should be higher in the case of complex technologies. Diffusion hampering effects of social proximity in the case of some technologies, however, comes at a surprise. An explanation for this might be found in firms’ aims in collaboration activities. Besides learning, the goal of a partnership can be the sharing of resources and costs (Kogut et al., 1992; Gulati, 1998). Technology adoption may cannibalize existing investments. This may decrease the potential net benefits of adaptation to a degree that firms decide to not adopt new technologies themselves. In these cases, firms may rather choose to rely on partners who have already adopted these technologies and may provide solutions based on these technologies at any time. Especially in the 19th century, firms faced large capital-intensive investments when establishing factories and infrastructure (Usselman, 1991). In these situations, further investments into new technologies may have been particularly unattractive. It is also risky, however, to ignore new technologies; hence, they collaborated with partners (in other regions) that had entered these technologies. In essence, this is an outcome of firm and regional specialization processes that are coordinated through collaboration (Hidalgo, 2015). If this strategy was not possible due to a lack of partners, firms had to adopt a new technology. While this explanation is speculative and we lack direct confirmation, it fits to the observed coefficients of social proximity and the overall patterns of firm/regional specialization. The meta-regression considering the complexity of technologies does not add any further insights to this, as the relationship between complexity and the effect strength of social proximity is insignificant (see Table 2.5). Consequently, we cannot confirm Hypothesis 3.

2.5 Conclusion

For nations and regions, mastering complex technologies is the path to long-lasting economic prosperity (Romer, 1990), as the capability to frequently innovate in the area of complex technologies leads to competitive advantages for regions (Porter, 1990). This creates products that are demanded beyond an individual region and will thereby add to economic growth (North, 1955; Kaldor, 1970) (Hidalgo et al., 2007). As much as we know about the benefits of being able to develop and apply complex technologies, however, we know much less about their generation and diffusion. In particular, the latter has received almost no attention in the literature thus far.

This paper seeks to shed some light on this issue. From previous studies it is known that regional characteristics and relational ones such as geographic, technological and social proximity facilitate the diffusion of knowledge in space (e.g., Hägerstrand, 1967; Jaffe et al., 1993; Bednarz and Broekel, 2019). We have built on this and add the consideration of technologies' complexity. More precisely, we elaborate how the strength of regional and relational factors in technologies' diffusion varies between simple and complex technologies.

To empirically assess this, a two-step procedure has been applied and US patent data of the 19th and early 20th centuries has been utilized. This time period was characterized by the start of modernity with inventions such as the railway and electricity. However, in comparison to today, there were still very limited communication possibilities, with personal communication and movements of inventors being primary channels of knowledge diffusion (Morrison et al., 2018).

The patent data allowed us to analyze the spatial diffusion process of 285 technologies over a time period of a hundred years. Empirically, the effect of proximities on the speed of technology adoption in space was evaluated with Bayesian diffusion models. Subsequently, the obtained coefficients of each variable in the diffusion models have been implemented in linear meta-regressions with technological complexity as the meta-independent variable. Thereby, we analyzed whether their effect strength is significantly related to the level of technological complexity.

The meta-analysis reveals that the level of technological complexity shapes the way technologies diffuse spatially. We find a positive relationship between complexity and the likelihood to observe contagious diffusion patterns. Thus, complex technologies tend to diffuse in wave-like patterns rather than in hierarchical forms. This finding is in line with Balland and Rigby (2017), who find geographic proximity to become more important with higher levels of technological complexity. Additionally, the meta-analysis shows that the effect strength of technological proximity, regions' population and technological diversity vary significantly between simple and complex technologies. In the case of complex technologies, experience with a diverse set of related technologies is even more beneficial than with simple technologies. This supports the idea that complex technologies require more prerequisites to be adopted quickly and successfully. This also strengthens the argumentation of Blaut (1977) and Ormrod (1990) of local contexts being important for understanding who does and who does not adopt innovations. Interestingly, the results of social proximity vary significantly between technologies, ranging from a positive to a negative relationship to adoption time. Moreover, no significant association with complexity can be observed. Therefore, we leave the elaboration of social proximity and technological diffusion to future research.

As it is typical for empirical studies, this chapter is subject to certain limitations. First, we focus on the early diffusion of technologies, i.e., mainly in the 19th century. Due to this focus, we are able to study the first one hundred years of diffusion for almost all technologies. Thereby, however, we fail to consider the most recent decades. These are characterized by the broad diffusion of the personal computer and the introduction of the internet. Both events may have significantly shaped the diffusion of technologies, as they have changed the way we work and communicate. However, we believe that the possibility of exploring the long-term diffusion of technologies in space fills a greater gap in the literature. This is particularly the case as a number of existing studies already focus on more recent years (see, e.g., Balland and Rigby (2017) and Sorenson et al. (2006)). Nevertheless, as a prospect for future research, it will certainly be fruitful to extend our research and empirical approach to more recent time periods.

Second, although we focus on the 19th century, we cannot rule out that some technologies existed before the start of our observation period in 1836. Thus, some technologies might already have diffused for a couple of years, whereas others newly emerge. We have partly tackled this issue by excluding all technologies that seem to emerge in 1836. Nevertheless, for future research to analyze the early emergence of technologies, it will be fruitful to extend our data to the year 1790.

Third, we have been restricted to studying the diffusion of technologies within the US, an example of Thompson's "national systems of cities" (1972). Most likely, our results will differ in other regions of the world such as Europe or Asia (Howells, 2002). For example, Europe is culturally more heterogenous and is divided by several language barriers. Therefore, it might be a fruitful region to analyze the effect of institutional proximity on technology diffusion. Low institutional proximity may especially hinder the diffusion of complex technologies, as communication and learning will be even more difficult in this case. Therefore, future research should consider other regions and evaluate our results in that context. In addition, future research should extend the national scope of analysis to an international one. Today, many firms operate (research) facilities in different countries to gain access to foreign technological knowledge not available in their home regions (Schaefer and Liefner, 2017). The activities of these multi-national firms might also shape the spatial diffusion patterns of technologies. For example, the German company Volkswagen operates "innovation hubs" in Germany, Israel, China and the US. Such locations might shape the international city hierarchy (Cantwell and Iammarino, 2000) and explain which cities adopt new technologies first, because it seems likely that employees of these firms regularly exchange information and knowledge.

To summarize, this study adds a more detailed picture to the diffusion of technology by combining the notions of complexity and proximity with spatial diffusion patterns. Nevertheless, our knowledge about the processes of spatial diffusion is still scarce and, thus, offers many opportunities for future research in economic geography.

Appendix

2.A1 Data description

Table A2.1: Descriptive Statistics of $CPLX_{ARVG}$ from original and refined data set

Data set	Variable	N	Mean	Median	St. Dev.	Min	Max
Original	$CPLX_{ARVG}$	436	3.90	3.64	1.28	1.19	11.21
Sample	$CPLX_{ARVG}$	285	3.78	3.57	1.17	1.59	9.19

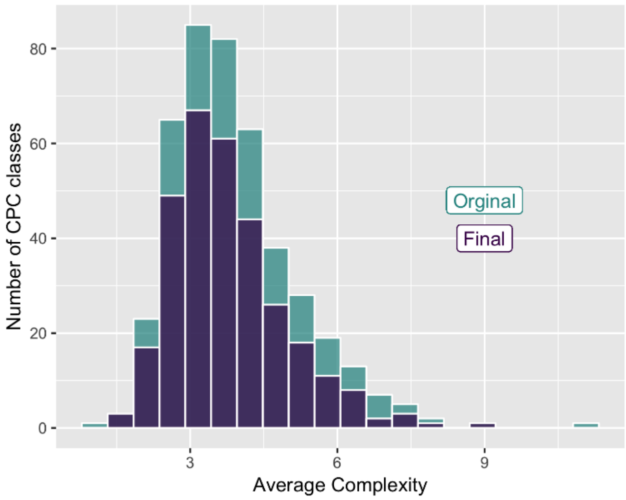

Figure A2.1: Stacked histogram of average complexity ($CPLX_{ARVG}$) for the original and refined data set

2.A2 Simultaneous diffusion "in all directions,"

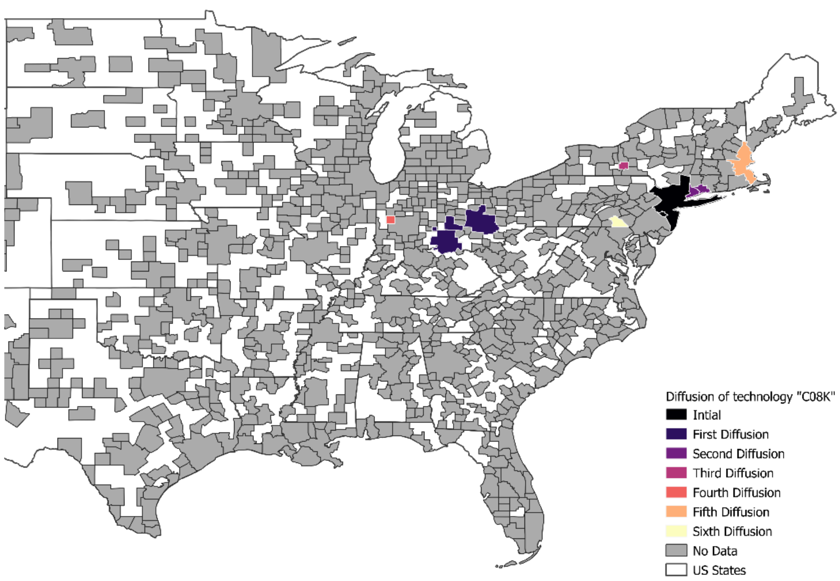

Figure A2.2: First 20 years of diffusion of technology "C08K" separated in to six diffusion steps.

The relationship of policy induced R&D networks and inter-regional knowledge diffusion

Abstract

Knowledge diffusion is argued to be strongly influenced by knowledge networks and spatial structures. However, empirical studies primarily apply an indirect approach in measuring their impact. Moreover, little is known about how policy can influence the spatial diffusion of knowledge. The chapter seeks to fill this gap by testing empirically the effects of policy induced knowledge networks on the propensity of inter-regional patent citations. We use patent citation data for 141 labor market regions in Germany between 2000 to 2009, which is merged with information on subsidized joint R&D projects. Based on the latter, we construct a network of subsidized R&D collaboration. Its impact on inter-regional patent citations is evaluated with binomial and negative binomial regression models. Our findings do not indicate that inter-regional network links created by public R&D subsidies facilitate patent citations and, hence, inter-regional knowledge diffusion.

This chapter is co-authored with Tom Broekel. The PhD candidate is the first author of the article. The paper has been published in the Journal of Evolutionary Economics, 2019, 29, pages1459–1481 https://doi.org/10.1007/s00191-019-00621-2

3.1 Introduction

To sustain innovation and gain competitive advantage, firms and regions need to innovate constantly, which requires the utilization of internal and external knowledge (Lundvall and Johnson, 1994; Sternberg, 2000). The utilization of external knowledge requires the diffusion of knowledge between organizations and across space. Since the pioneering work of Hägerstrand (1967) it is known that knowledge does not diffuse frictionlessly within socio-economic systems. Despite technological advantages, geographic distance remains a significant obstacle. Other factors such as cognitive, institutional, organizational, and social distances add to an unequal, selective, and potentially too low diffusion of knowledge (Buisseret et al., 1995; Boschma, 2005). Policy has recognized this and established measures seeking to stimulate knowledge diffusion, foremost by subsidizing joint R&D projects (Buisseret et al., 1995; Broekel and Graf, 2012). The EU-Framework Programmes are a well-known example for such an initiative (Scherngell and Barber, 2009).

Despite such initiatives existing for decades, evidence for a positive contribution to knowledge diffusion remains scarce. While their effects on firms' and regions' innovation activities has been frequently documented (Hewitt-Dundas and Roper, 2010; M. A. Maggioni et al., 2014; Broekel, 2015), little to no evidence exists on their effectiveness for inter-organizational or inter-regional knowledge diffusion. The present chapter seeks to fill this gap by using an empirical approach frequently applied to study the diffusion of knowledge in space. Following the work of (Jaffe et al., 1993), we used patent citations as indications of inter-regional knowledge diffusion and tested for the effect of joint R&D projects subsidized by the German Federal Government. Hence, besides modelling knowledge diffusion directly, we contribute to the literature by studying the effect of subsidized knowledge networks among regions of a single country, as most existing studies focus on knowledge diffusion at the European level (Maggioni et al., 2011).

In the empirical analysis, we use information on patents and patent citations for 141 German labor market regions between 2000 and 2009. On this basis, technology-specific knowledge diffusion, regions technological relatedness and co-inventor relations are established. Following Broekel and Graf (2012), we construct policy induced networks emerging from the subsidization of joint R&D projects by the German federal government, which are matched to the patent data.

The chapter is structured as follows. The second section discusses the mechanism of knowledge diffusion between organizations and in space. The empirical data and variables are described in section 3.3. Section 3.4 introduces the empirical approach. The results of the analyses are presented in section 3.5. Section 3.6 concludes the chapter with a short summary, and discussion on its limitations, future research prospects and political implications.

3.2 Theoretical considerations

3.2.1 Knowledge diffusion, networks, and proximities

It has long been established that knowledge creation and innovation are closely linked to knowledge diffusion. Kline and Rosenberg (1986) emphasize the central role of frequent knowledge exchange between economic actors in their "chain-linked model". Here, innovation processes consist of feedback loops, forward and backward linkages, as well as other interactions between actors internal and external to an organization. In line with this, Schrader (1991) confirms a positive correlation between firms' economic performance and the frequency of informal inter-firm interactions. Zucker et al. (1998) add to this by finding cooperation between biotechnology firms and scientists positively impacting product innovation.

In general, knowledge diffusion can be described by individuals (senders and recipients) (un)intendedly sharing their knowledge via (in)direct communication (Witt et al., 2012). Successful knowledge diffusion dependents on the senders' intentions and capacities to communicate (Ibid.) as well as on the recipients' willingness and abilities to recognize and to absorb knowledge (Cohen and Levinthal, 1990; B. Lundvall and Johnson, 1994). The ability to assimilate new knowledge tends to be greater "when the object of learning is related to what is already known" (Cohen and Levinthal, 1990: p. 131). Therefore, it is positively influenced by the cognitive overlap of two actors' cognitive "interpretation system[s]" (Nooteboom et al., 2007: p. 1017). These arguments are frequently summarized as cognitive (alternatively technological) proximity between actors positively influencing the likelihood and effectiveness of knowledge diffusion. In addition to cognitive proximity, Boschma (2005) argues in favor of four additional proximities that make knowledge diffusion more likely and effective: organizational, institutional, social and geographic proximity. Organizational proximity is a control-related dimension: relations with high organizational proximity are embedded in the same business routines, hierarchies, and value systems, e.g., being part of the same firm (Loasby, 2001; Balland, 2012). Such circumstances tend to help dealing with uncertainty and opportunism, which in turn facilitates knowledge diffusion (Boschma, 2005). Similar conclusion can be drawn when actors are embedded in the same institutional framework, which implies that they share the same norms and formal rules (Ibid.).

(Hägerstrand, 1952; 1966; 1967) promotes the view of geographic distance being an (and maybe "the") explanatory factor for knowledge diffusion. Geographic proximity tends to facilitate face-to-face interactions and mutual trust, which are necessary conditions for successful knowledge exchanges (Howells, 2002). Boschma (2005) points out that geographical proximity is neither necessary nor sufficient for knowledge diffusion, however, as other forms of proximity are similarly crucial. Besides geographic proximity, Hägerstrand (1965) also suggests knowledge diffusion being closely related to the embeddedness of actors into social relations, e.g. kinship or friendship. Social interactions consist of multilateral, interdependent and multilevel network relations that may serve as (in)direct knowledge channels (Tijssen, 1998; Maggioni et al., 2007). This is commonly referred to as social proximity (Boschma, 2005).

3.2.2 Knowledge diffusion and proximities – what about R&D policy?

Given that one or multiple of these proximities are frequently absence or weakly developed, inter-organizational knowledge diffusion can be expected to be below a social optimum. As this may reduce innovation, policy intervention can be justified (Buisseret et al., 1995). While different approaches exist to deal with this, policy mostly provides monetary incentives in the form of subsidizing joint R&D projects to increase inter-organizational learning and knowledge diffusion (Breschi and Cusmano, 2004; Broekel and Graf, 2012; Broekel et al., 2015). Usually, organizations can apply for project funding within the scope of policy-defined calls and by providing information on their projects' aims, required resources, partners, and expected outcomes. A well-known example of such an initiative is the EU-Framework Programme (EU-FRP). The EU-FRP has become the most important R&D policy tool of the European Union. For the period 2014 to 2020 it involves about 80 billion €, which are granted in the form of subsidies to R&D projects. For a more detailed discussion see e.g. Breschi and Cusmano (2004). Here, the subsidies are exclusively granted to projects in which organizations conduct R&D in a collaborative manner. Similar programs also exist at the national level. For instance, in Germany, the Federal Ministry of Education and Research as well as the Federal Ministry of Economics and Energy invest about 3-4 billion € each year as subsidies for project-based R&D. About 30 % of the subsidies are granted to collaborative R&D projects (Broekel and Graf, 2012). Crucially, cooperating organizations are obligated to write a cooperation agreement in which they grant access to their intellectually property rights that are necessary to conduct the project. In these cases, financial support is only granted when all participants agree on exchanging knowledge within the scope of the project. This includes the use of property rights, access to technical expertise and regular face-to-face meetings (BMBF, 2008). Accordingly, subsidized joint R&D projects bear the potential of facilitating if not enabling inter-organizational knowledge diffusion (Broekel and Graf, 2012). The rather limited existing research on this issue shows that knowledge generation of regions and organizations is positively associated with participating in subsidized joint R&D projects. For example, (Fornahl et al., 2011) discover that biotechnology firms engaged in subsidized collaborative R&D have higher patent output than firms not receiving such funds. Maggioni et al. (2014) and Broekel (2015) confirm this for the regional level. However, we argue that these studies do not provide direct empirical evidence for a direct knowledge diffusion-enhancing effect of subsidized joint R&D projects.

3.2.3 The indirect and direct approaches of analyzing spatial knowledge diffusion

Traditionally, knowledge diffusion studies have investigated the inter-organizational[4] knowledge diffusion by empirically tracking inventions, applications, or products over time and space (see e.g. Hägerstrand, 1952; 1965; 1967; Rogers, 2003). The primary interest of these studies is to find re-occurring patterns of (spatial) knowledge diffusion and to analyze the extent to which these dimensions represent significant obstacles to knowledge diffusion. More recently, this research tradition, among others, has stimulated the analysis of publication and patent citation patterns, which are seen as indications of knowledge diffusion (Jaffe et al., 1993). By studying the spatial structure of patent citations, these authors show that patens are more likely to cite other patents if their inventors are located in geographic proximity. Breschi and Lissoni (2009) extend this research by using the same approach and identify limited geographic mobility of inventors being primarily responsible for this finding, as they tend to cite their older patents assigned to their previous employers. Since then, the evaluation of patent (and publication) citation has been used frequently to evaluate spatial knowledge diffusion (Peri, 2005; Maggioni et al., 2007; Hoekman et al., 2009; Paci and Usai, 2009). Interestingly, this, what we call, the "direct approach", has not been used to study the effectiveness to subsidized joint R&D projects, as vehicles of knowledge diffusion.

When testing the effect of subsidies for R&D projects on inter-organizational knowledge diffusions, researchers seem to be primarily inspired by another literature that focuses on the relevance of spatial knowledge spillovers. In this line of research, studies relate the innovation output of individuals, organizations, and regions to the knowledge potentially available to them (e.g. Bode, 2004). Inspired by the classic knowledge production function approach, Griliches (1979) and Jaffe (1986) argue that the difference between observed innovation output and

[4] For the sake of readability, we exclusively use the term inter-organizational knowledge diffusion. However, the arguments apply in an identical fashion to knowledge diffusion between individuals or regions.

knowledge of an innovation-generating entity (inventors, organizations, regions) can partly be explained by the entities use of external knowledge that was absorbed through various mechanisms, i.e. by knowledge diffusing between the entities. In the spatial knowledge spillover literature, researchers extend the individual knowledge input of entities by their potential access to external knowledge. For instance, studies assessing the innovation performance of regions consider a wide range of regional characteristics approximating their knowledge endowment, e.g. presence of high-tech industries, universities, etc. (Broekel and Brenner, 2011). This set of factors is extended by variables describing the knowledge endowment of neighboring regions (usually the spatial lag of their innovation output). If the latter show a positive correlation with the innovation output of the focal region (while controlling for its own knowledge endowment), it will be interpreted as a confirmation of knowledge having diffused from the neighboring regions into the focal region and thereby having contributed to its innovation generation (Bottazzi and Peri, 2003).

The same approach can be transferred to test for the effect of knowledge networks. In this case, the positions of entities in inter-organizational knowledge networks are plugged into the knowledge production function approach replacing or complementing the spatial lag of other entities' innovation output (Maggioni et al., 2014). The latter is usually described in terms of their local (degree) or global (betweenness / eigenvector) centrality. Larger centralities imply a "better" access to the knowledge potentially diffusing in the network. Hence, if the innovation output (or its change) is found to be statistically related to the centrality measures, it is inferred that the correlation is caused by the better access and ultimately use of the diffusing knowledge.

The before mentioned studies by Fornahl et al. (2011) and Maggioni et al. (2014) use this approach to assess the significance of subsidized knowledge networks for knowledge diffusion. While this "indirect approach" is elegant and has its merits, it has one major flaw in this context: whether knowledge is actually diffusing between entities and is utilized, remains unobserved. This implies that the evidence for an enhancing effect of R&D subsidies on knowledge diffusion, which is largely based on this approach, remains subject to interpretation.

The present study therefore seeks to complement the existing evidence on the effectiveness of R&D subsidies using the "direct approach". Following Jaffe et al. (1993), we use patent citations to approximate knowledge diffusion. While it is also troubled by the unobserved knowledge sourcing and utilization, it is a much more direct approach and avoids many issues of spurious correlation that are likely troubling the indirect approach.

3.3 Data and empirical approach

3.3.1 Modelling knowledge diffusion

We follow Jaffe et al. (1993) and Breschi and Lissoni (2009) and rely on patent citations to model knowledge diffusion. Within the knowledge spillover literature, patent citations are an often used “paper trail” that “track” knowledge diffusion (e.g. Jaffe et al., 1993; Peri, 2005). Patent citations are argued to be an indicator of knowledge transfer and accumulation, as the citing patent builds its knowledge upon a piece of the existing knowledge of the referred patent (Jaffe et al., 1993). According to Howells (2002), patents are a proxy for tacit knowledge flows, since it is necessary to recognize, understand, and recombine the codified knowledge imprinted in the patent.

Making use of patent citations, however, brings along some issues and limitations that must be carefully considered (Breschi and Lissoni, 2005). First, patent citations may only account for knowledge flows that are encodable, commercially exploitable and legally patentable (Criscuolo and Verspagen, 2008). Moreover, the propensity to patent in the first place significantly differs between industries; e.g. only 8% of product innovations are patented in the textile industry, while the share is almost 80% in pharmaceuticals (Arundel and Kabla, 1998). Therefore, patents capture only a portion of the created knowledge and correspondingly, their citations only a fraction of the knowledge circulating between innovators (Criscuolo and Verspagen, 2008). Second, patent citations may be a noisy indicator as usually not all citations are made by the inventor himself. Submitted patents will be checked by examiners who probably add further citations where appropriate (Jaffe et al., 1993). Fortunately, the European Patent Office (EPO) specifies which citations have been made by the inventor and which have been added by the examiner. Moreover, it has been argued that, at the regional level (which we will use in the later analysis), these issues are of smaller relevance and patent citations may be a good indicator of possible spillovers, even though it is not sure that these potentials have been realized in all cases (Paci and Usai, 2009).

The data on patents and citations is taken from the REGPAT and Citations Database provided by the OECD. In contrast to other patent data sources, patents have already been assigned to the regions of inventors’ residence. Patents that have been developed by more than one inventor have been weighted accordingly (fractional counting). Considering the beforementioned limitations, we exclusively rely on patent citations identified as ‘inventor citations’. Moreover, to diminish biases stemming from intra-organizational citations or inventors’ self-citations, we remove all intra-regional citations. We also remove citations from different inventors and regions that share the same applicant (company) and so will further reduce the likelihood of theses citations representing intra-organizational knowledge flows.[5] To reduce the biases stemming from variations in industries' patent and citation activities, we restrict the analysis to individual technologies, i.e. we only consider technology-specific knowledge flows, that is, we exclusively consider citations between patents belonging to the same IPC-subclass (4-digit level).

The empirical analysis is conducted at the regional level because we are interested in spatial knowledge diffusion (e.g. Maggioni et al., 2014; Broekel, 2015). The approach avoids potential biases caused by unclear assignments of patents to applicants. We use the 141 German labor market regions classified by Kosfeld and Werner (2012). Labor market regions are an often-applied regional unit of analysis (e.g. Frenken et al., 2007; Broekel et al., 2015), especially when patent data are used. They ensure that an inventor 's workplace and residence are most likely located in the same region and, hence, in the same unit of analysis (Broekel, 2015).

The most important variable is the citation frequency between two regions. The variable, denoted as $CIT_{i,j,t,s}$, is constructed as follows. In year *t*, a knowledge flow exists from region *i* to *j* when at least one inventor living in region *i* cites at least one patent assigned to an inventor in region *j* within technology *s*. See exemplary Figure 3.1A for the patent citations in technology class H04L. More precise, $CIT_{i,j,t,s}$ is the count of citations from region *i* to *j*. We employ a five-year moving window for the patent and citation counts to control for yearly fluctuations. For example, the moving window of region i's inventors citing region j's inventors in 2005 includes citations made from 2005 to 2009. The five-year moving window is consistent with the literature since most patents lose their economic impact within this time frame (Griliches, 1979).

3.3.2 Knowledge diffusion channels and regional characteristics

The variable, in which we are most interested is the intensity of subsidized R&D projects, in that organizations of two regions jointly participate. Data on subsidized collaborative R&D projects is obtained from the subsidies catalogue of the German federal government ("Förderkatalog"). According to Broekel and Graf (2012) and Broekel (2015), joint R&D projects listed in this database are intended to stimulate collective learning and knowledge diffusion. For instance, to acquire financial support for a joint R&D project, all participants must agree on certain rules that facilitate collaboration and knowledge exchange. Moreover, all intellectual property rights lying within the scope of the project and that existed before project

[5] Note that there is no direct assignment of applicants to inventors. This is particularly true, when multiple inventors are given on an applicant's patents.

start must be revealed if demanded by the partners (BMBF, 2008). The partners moreover have frequent face-to-face interactions in consortia meetings. They usually have a proven track record of R&D and have convinced application reviewers that their projects (and their partnership) have the potential to become successful.

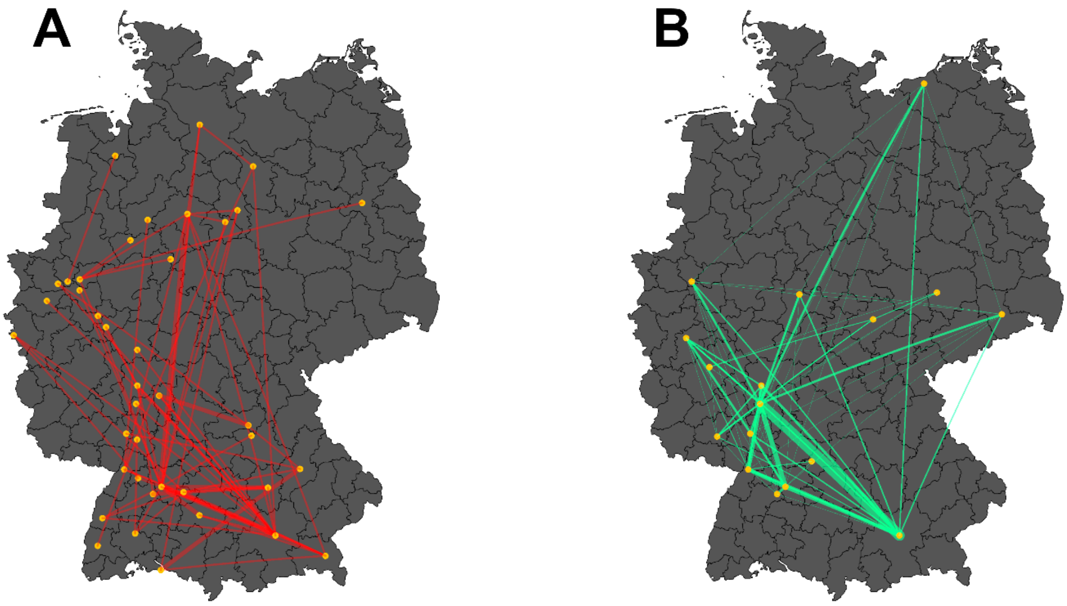

Figure 3.1:Spatial networks of technology H04L "Transmission of digital information" **A** Inter-regional patent citations in 2003 and **B** Subsidized R&D network in 2000

All projects listed in the subsidies catalogue have been assigned to a "Leistungsplansystematik" (LPS), which is a thematic classification scheme similar to the IPC. The classification includes 22 main classes, for example, biotechnology or nanotechnology. These main classes are further divided into more differentiated subclasses (e.g. photonics (class: I25020)) (Broekel and Graf, 2012). Based on these subclasses and the project descriptions, we manually matched 87 four-digit IPC subclasses to about four hundred thematic subclasses of the subsidies catalog. Accordingly, we can differentiate between 87 technologies for which information on R&D subsidies and patenting are available. In line with Broekel and Graf (2012), we constructed the variables $SUBS.NET_{i,j,(t-3),s}$ and $SUBS.NET_{i,j,(t-5),s}$ as follows: a direct subsidized link between two regions i and j exists if at least two of their organizations participated in the same R&D project in year t (see exemplary Figure 3.1B for the R&D network of technology class H04L). To model the intensity of their joint R&D work, the variable $SUBS.NET_{i,j,t}$ captures the count of the regions co-involvement in subsidized R&D projects. To avoid biases in the coding of joint projects, we restricted the analysis to projects with at least three participants. Moreover, we neglected all intra-regional relations as these are of no interest in the context of the chapter.

In accordance to the patent data, the subsidies data include information on the projects' exact starting and ending dates making the construction of a moving window unnecessary. We consider a three- and five-year time lag to the citation variable $CIT_{i,j,t,s}$ as we suspect a time lag given between the cooperation, i.e. the potential knowledge exchange, and the patent citation, i.e. the usage of this knowledge. To make sure that our results are robust, we decide to test for two time lags.

Different types of proximity influence the likelihood of knowledge diffusion between two regions (see Section 4.2.2) [6]. We mainly included these into the models as control variables. The first considered knowledge diffusion channel approximates social proximity. The intensity of social proximity is measured by the strength of co-inventor ties. The co-inventor network has been constructed in two versions denoted as $CO.INV_{i,j,(t-3),s}$ and $CO.INV_{i,j,(t-5),s}$. A co-inventor relation between regions i and j exist when two inventors, one from each region, have worked together on the same patent in technology s. It is usually interpreted as these two inventors having met personally and hence having established a personal relationship. Accordingly, the frequency of such co-inventor relations signals the strength of the social relations (of inventors) between two regions. We included this variable in a three and alternatively a five-year time lag to the citation variable, to avoid problems of endogeneity and the possibility that citations directly emerge from co-invented patents.

The other proximity types are defined as follows. Geographic proximity is captured by the variable $DIST_{i,j}$, which is specified by the great circle distance between the centroids of region i and j in kilometers and divided by one hundred in order to scale the variable.

The construction of technological proximity ($TECH_{i,j}$) is somewhat more complex. We follow the early work of Jaffe (1986), Engelsman and van Raan (1994) and Verspagen (1997) and estimate the cosine similarity between regions' technological profiles. In practice, for each of the 141 German labor market regions, we count the applied patents of each IPC subclass resulting in a 141 (regions) x 629 (technologies) matrix. In the next step, we calculated the cosine similarity between each region and obtain the 141x141 technological relatedness matrix for 2000 to 2009, which serves as a technological proximity measure.

Boschma (2005) describes organizational proximity as a control-related dimension: high organizational proximity means that relations are embedded in the same organizational routines

[6] Note that the data at hand does not allow to capture institutional proximity.

and hierarchies. We sought to capture this by the organizational proximity ($ORG_{i,j}$) measure. It represents the number of patents that are invented by the same organization with inventors from both regions *i* and *j*.

Table 3.6: Descriptive statistics

Variables	Mean	SD	Median	Min	Max
$CIT_{i,j,t,s}$	0.226	0.816	0	0	28
$CO.INV_{i,j, (t-3),s}$	0.207	1.771	0	0	81
$CO.INV_{i,j, (t-5),s}$	0.23	1.668	0	0	84
$SUBS.NET_{i,j,(t-3),s}$	0.014	0.351	0	0	47
$SUBS.NET_{i,j,(t-5),s}$	0.012	0.282	0	0	45
$DIST_{i,j}$	3.56	1.86	3.43	0.16	9.52
$TECH_{i,j}$	0.410	0.173	0.41	0	0.92
$ORG_{i,j}$	0.007	0.03	0	0	6
$PAT_{i,t}$	6.846	15.01	0	0	207.1
$PAT_{j,t}$	7.159	15.75	0	0	207.1
$SUBS_{i,t}$	0.571	1.305	0	0	46
$SUBS_{j,t}$	0.597	1.268	0	0	46
$CIT.LAG_{i,j,t}$	0.109	0.586	0	0	28

The likelihood of inter-regional knowledge flows is also dependent on regions' individual characteristics. The first variable accounts for regions' innovative output ($PAT_{i,t,s}$ and $PAT_{j,t,s}$), i.e. the sum of technology specific patents created by inventors (factual counting) located in the respective regions *i* or *j* in time period *t* and technology *s*. The more a region patents, the more likely it will receive citations and cite other regions' patents.

Moreover, we consider the amount of subsidies that regional actors acquired in period *t*, ($SUBS_{i,t,s}$ and $SUBS_{j,t,s}$). More precisely, $SUBS_{i,t,s}$ and $SUBS_{j,t,s}$ are the sum of granted subsidized projects to organizations located in region *i* and *j* in technology *s* and year *t*. $SUBS_{i,t,s}$ and $SUBS_{j,t,s}$ are included for two reasons; first, large sums of subsidies may indicate the presence of extensive R&D activities, which in turn make citations more likely and second, while a significant coefficient of *SUBS.NET* is the most direct confirmation of R&D subsidies enhancing inter-regional knowledge diffusion, the coefficients of *SUBS* can still give some indication as to whether the subsidization of R&D activities will lead to larger citation activities or to research that is more frequently cited.

Descriptive statistics for all variables are listed in Table 3.1.

3.3.3 Empirical modelling

Our aim is to explain the intensity of knowledge flows ($CIT_{i,j,t,s,}$) between regions with the variables introduced above and particular emphasis on the impact of R&D subsidies. In a common manner, we employed a gravity model approach. Based on the work of Isard (1954) and Tinbergen (1962), the gravity approach is a conceptual framework frequently used in investigations of trade and migration flows as well as tourism and commuting interactions (Burger et al., 2009). It has also previously been used to estimate knowledge flows between regions (see Peri, 2005; Maggioni et al., 2007; Hoekman et al., 2009; Paci and Usai, 2009).

The basic expression of a gravity model can be written as follows (Burger et al., 2009; Hoekman et al., 2009):

$$I_{i,j} = \frac{K^{\beta_1} MASS_i^{\beta_2} MASS_j^{\beta_3}}{DISTANCE_{i,j}^{\beta_4}}$$

where $I_{i,j}$ is the intensity of interaction between two entities, i.e. the amount of knowledge flows between regions *i* and *j*. The "masses" of origin *i* and destination *j* are individual regional characteristics potentially influencing this interaction intensity. In this chapter, the number of regional patents (*PAT)* and subsidies *(SUB)* represent the "MASS" variables.

$DISTANCE_{i,j}$ represents the dyadic relations (e.g., geographical distance) between region *i* and *j*. Crucially, the model is not restricted to two masses and one distance variable, as in practice; it is logarithmized, transforming the equation into a standard (log)linear regression equation, which can include many more explanatory variables (see for a detailed discussion Broekel et al. (2014)).

Our data sample consists of patent citation relations among 141 German labor market regions in 87 IPC classes over a period of ten years (2000-2009). This gives us slightly more than 16 million observations. 34,487 of these are positive ($CIT_{i,j,t,s} > 0$). Notably, we defined a positive case if at least one citation occurred between a pair of regions in any of the ten years. Fitting a regression model to this vast amount of data with just 0.2% non-zero observations would lead to methodological as well as computational difficulties. We therefore reduce the data by randomly matching one positive case with two negative cases. More precise, each positive case of $CIT_{i,j,t,s} > 0$ is matched with two zero-cases of $CIT_{i,j,t,s} = 0$, whereby one of these zero-cases is randomly drawn from the zero-cases of the citing region *i* and one from the set of zero-cases of the cited region *j*. Hence, our zero-cases share many characteristics of the positive cases.

The matched sampling has 255,493 observations of which 34487 are positive ($CIT_{i,j,t,s} > 0$), i.e. 13,5% of the sample. The ratio is less than 1:3 because the figure is estimated considering all years, while $CIT_{i,j,t,s} > 0$ might only be positive in one.

Given that our dependent variable, $CIT_{i,j,t,s}$, is a count variable, Poisson or negative binomial regression models are generally appropriate. We decided to calculate a binomial and a negative binomial regression. The binomial regression will provide insights into the probability of inventors in two regions citing each other at all. The negative binomial part seeks to explain the variance in the citation intensities for the sample of observations with at least one citation between 2000 and 2009.

The regressions utilize the panel structure of our data, which allows for including technology- and time-fixed effects controlling for unobserved time invariant effects. Moreover, we also included the lagged version of our dependent variable ($CIT.LAG_{i,j,t,s}$) to capture any remaining unobserved and time-variant structures. It is lagged by six years to the dependent variable ensuring no overlap in the citations used in its construction.

3.4 Empirical results

We estimated four regression models in total: two binomial models distinguishing between inter-regional citations and no inter-regional citations and two negative binomial models explaining the magnitude of inter-regional citations, both with varying time lags of three and five years. The results are presented in Table 3.2 (three-year time) and Table A3.1 in the Appendix (five-year time lag). The results are very robust across all models. We will therefore interpret all outcomes at once.

Mostly, our control variables behave according to our expectations. Increasing geographical distance ($DIST_{i,j}$) between two regions decreases the probability of positive citation counts (binominal model) and the magnitude of inter-regional citations (count model). This confirms the geographically localized nature of knowledge spillover (Jaffe et al., 1993; Storper and Venables, 2004). Technological proximity ($TECH_{i,j}$) has a positive and significant impact on the existence of inter-regional citation in general (binominal model) and on their intensity (count model). The finding is in line with those of Peri (2005) and supports the idea of actors being more likely to seek and absorb new knowledge in cognitive proximity to what they already know (Cohen and Levinthal, 1990). Contrary to our expectations, organizational proximity ($ORG_{i,j,t}$) mostly obtains a significantly negative coefficient in the binomial model. It means that region pairs with strong organizational linkages are less likely to cite each other. We suspect that this finding might be caused by our data cleaning in which we removed all citations associated with patents of the same applicant and with the inventors residing in two or more regions. Accordingly, in the construction of the dependent variable, we removed intra-organizational citations, which are closely related to patents with multiple inventors and one applicant, which provide the underlying information for $ORG_{i,j,t}$. In light of this, it is somewhat surprising to see $ORG_{i,j,t}$ gaining a significantly positive coefficient in the count model, suggesting that once citation links are established between regions, organizational proximity facilitates these and supports knowledge diffusion.

Table 3.7: Results of the fixed-effects, binomial and negative binomial (count) regression models (3-year time lag)

	Binomial model	**Count model**
Variables	**Estimate (SE)**	**Estimate (SE)**
Policy		
$SUBS.NET_{i,j,(t-3),s}$	0.011	-0.007
	(0.013)	(0.007)
$SUBS_{i,t,s}$	0.013***	0.005***
	(0.003)	(0.002)
$SUBS_{j,t,s}$	0.021***	0.002
	(0.003)	(0.002)
Proximities		
$CO\_INV_{i,j,(t-3),s}$	0.109***	0.012***
	(0.005)	(0.001)
$DIST_{i,j}$	-0.098***	-0.017***
	(0.004)	(0.003)
$TECH_{i,j}$	1.816***	0.124***
	(0.038)	(0.027)
$ORG_{i,j}$	-0.015***	0.005***
	(0.004)	(0.001)
Regional properties		
$PAT_{i,t,s}$	0.017***	0.004***
	(0.0005)	(0.0003)
$PAT_{j,t,s}$	0.014***	0.003***
	(0.0005)	(0.0002)
Control		
$CIT\_lag_{i,j,t,}$	-0.231***	0.046***
	(0.011)	(0.003)
Year Dummies	Yes	Yes
Technology Dummies	Yes	Yes
AIC	163,794	102,695
2x Log-likelihood		-102,526
Observations	255,493	34,487
Non-zero obs	34,487	-

* significant at the 90% level, ** significant at 95% level, *** significant at 99 % level

We interpret this contradictory finding as an indication of large organizations acting as gatekeeper (Giuliani and Bell, 2005). Sometimes they help other organizations in establishing inter-regional linkages, which they, for various reasons, are not able to build otherwise. In network science, this corresponds to the triadic closure argument: partners are likely to connect to partners of their partner (e.g., Granovetter, 1973). Accordingly, large organizations with facilities in multiple regions may stimulate other actors in these regions to interact with each other. Our results, however, suggest that this effect only comes into play when at least one organization other than the large one has managed to establish such a link. As $ORG_{i,j,t}$ is primarily included as a control variable, we refrained from exploring this issue in greater detail and leave it to future research.

The findings for our measure of social proximity clearly confirms the crucial role this plays for spatial knowledge diffusion. $CO.INV_{i,j,(t-3),s}$ and $CO.INV_{i,j,(t-5),s}$ are both significantly positive in the binomial and count model. Hence, as (Hägerstrand, 1966) proclaimed social networks are crucial for diffusing information about new knowledge and for subsequently spreading this knowledge.

The results for the individual patenting activities of citing ($PAT_{i,t,s}$) and cited regions ($PAT_{j,t,s}$) are highly significant and have a positive relationship with citations in the binomial and count model. Thus, confirming a size effect, regions with a larger patent output have a higher likelihood to cite and to be cited.

The last control variable, $CIT.LAG_{i,j,t,s}$ measures the citation intensity between regions in the previous period. It is significantly negative in the binomial model and significantly positive in the count model. Accordingly, if inventors in region *i* cite patents of inventors in region *j*, it is unlikely that they cite each other again in the subsequent period. However, in case they do, their citation intensity will grow. With some speculation, we can interpret this as a kind of “probing” behavior. Inventors will tap into other regions’ knowledge bases to solve specific problems. However, if the problem is solved, their interest in keeping these relations vanishes. In some cases, though, the regions’ knowledge bases turn out to be complementary, leading to increasing cross-citations and intensifying knowledge exchanges. The observation might also be related to the vary erratic patenting numbers characterizing many (and particularly smaller) regions (Burger et al., 2009), which translate into erratic citation numbers. In the case of regions with few patents, a decrease in these likely translates into the vanishing of many inter-regional linkages created by previous patents’ citations. While the decrease in patents will also lower inter-regional citation frequencies of regions with many patents, their cited patents’ set of regions is likely to remain intact due to the higher levels of citation frequencies.

Our results clearly support the idea of proximities and general economic as well as technological structures shaping the spatial diffusion of knowledge. The question remains, if policy can impact the diffusion by subsidizing R&D projects, which provide a framework for inter-organizational learning.

The variables $SUBS.NET_{i,j,(t-3),s}$ as well as $SUBS_{i,t,s}$ and $SUBS_{j,t,s}$ answer this question. $SUBS.NET_{i,j,(t-3),s}$ (and $SUBS.NET_{i,j,(t-5),s}$) remain insignificant in all models. Accordingly, joint projects with participants form multiple regions do not facilitate inter-regional knowledge flows, at least when these are measured by patent citations three or five years after the project. This clearly contradicts our expectations and contrasts with the findings of (Fornahl et al., 2011) and Broekel (2015), which, however, employ an indirect empirical approach. They investigate the relation between organizations and regions probable exposure to knowledge flowing through subsidized joint R&D projects and their innovation output. With the more direct approach used in this chapter, we fail to replicate their findings. There are multiple reasons that may cause this result. First, partners in subsidized joint R&D projects may use these subsidies as windfall gains by using subsidies to strengthen already existing collaboration, which knowledge exploitation potential has already depleted. Second, many subsidized R&D projects may also not result in any patents and, accordingly, citations. Third, subsidized joint projects may fail in establishing relationships that outlast the length of the project. While the knowledge exchange might happen during the project, the inventive process might exceed projects' durations (which is, on average, close to 36 months) and our time lags of three and five years. Fourth, the collaborative feature of joint R&D projects covered by our data may simply be insufficient for significant knowledge exchange. Clearly, further work, which may have more detailed data available, may shed additional light on these issues.

More promising results (from a policy perspective) are obtained for the number of subsidized projects acquired by regional organizations. The variables $SUBS_{i,t,s}$ and $SUBS_{j,t,s}$ are significantly positive in all binominal models, i.e., the more projects acquired by regional organizations, the more likely are these regions' inventors citing and getting cited. However, this is not consistently true in the count models. Here, only $SUBS_{i,t,s}$, is positively significant, i.e., the number of subsidized projects in the citing region. The number of projects acquired by the cited region ($SUBS_{j,t,s}$) is mostly insignificant. Accordingly, organizations in regions that are more successful in project acquisition are more likely to create (patent) output that refers to the work of inventors outside their region. Hence, the subsidization of projects seems to stimulate external knowledge sourcing. Unfortunately, whether this sourcing goes beyond already existing contacts cannot be tested. The insignificance of $SUBS_{j,t,s}$ implies that (subsidized) research of cited regions does not produce findings that are recognized and utilized by actors outside the region at an above average rate. Put differently, if citations are interpreted as an impact indicator, their research's impact does not seem to exceed the average. Hence, subsidies are granted to organizations in regions, which research impact is not outstanding.[7] This is in line with Broekel et al. (2015) who discovered that, in contrast to the EU-Framework Programmes, research excellence at the regional level does not seem to be the primary allocation determinant of R&D subsidies granted by the German federal government.

3.5 Conclusion

Despite being one of its central aims, few attempts have been made at evaluating the contribution of subsidies for joint R&D projects on inter-organizational and inter-regional knowledge diffusion. Most existing studies follow a knowledge production function approach to test for such policy's effectiveness (Maggioni et al., 2014; Broekel, 2015; Broekel et al., 2015). We argued that the traditional literature on knowledge diffusion offers a more direct approach for assessing the contribution of R&D subsidies to inter-regional knowledge diffusion. By following the work of Jaffe et al. (1993) and others, we used patent citations as indication of knowledge diffusion, which are related to subsidies for joint R&D projects by the German national government. We aggregated the data to the level of German labor market regions during the years 2000 to 2009. While our results confirmed the diffusion hampering effects of different kinds of distances (cognitive, geographical, social, organizational), we did not find clear evidence of subsidies significantly stimulating inter-regional knowledge diffusion. While they seem to enhance organizations' ability to source knowledge in other regions, subsidizing inter-regional collaboration does not increase the intensity of subsequent patent citation intensity. Accordingly, we fail to confirm existing evidence obtained by the indirect, knowledge production function type approach (Ponds et al., 2010; Fornahl et al., 2011; Broekel, 2015).

While our results fail to deliver support for project-based R&D subsidies achieving one of their primary objectives, there are a number of shortcomings that put our results into perspective. These may also serve as possible starting points for future research. First, by aggregating the data at the regional level, we reduced some of the inherent limitations of patent data. However, when data are available, the organizational level is certainly more appropriate

[7] Note that one must be very careful with this interpretation as it is prone to an ecological fallacy trap. We do not say (and find) that the research of organizations receiving R&D subsidies is "just" average or not excellent. Our results exclusively refer to the regional level.

for such kind of analyses. Consequently, the findings must be interpreted with caution, as there is the threat of aggregation biases or “ecological fallacy” (Downs and Mohr, 1976: p. 707). Second, while providing more general results, this and similar studies also lack the detailed insights, which can be obtained with qualitative research. For example, the “innovation biographies” approach by Butzin (2009) may be used to study the impact of R&D subsidies on knowledge diffusion in more depth. Third, the study also shares the well-known and frequently discussed limitations of patent data (Griliches, 1998). Particularly, patent citations may only account for knowledge flows that are encodable, commercially exploitable and legally patentable (Criscuolo and Verspagen, 2008). Thus, the chapter only considered knowledge flows that led to new patents and cite the corresponding ones. An organization might develop a new product, however, without patenting it, as the organization uses other ways of securing their intellectual property (Cohen et al., 2000). In this case, knowledge diffuses from one region to another without leaving a trail in patent data. The extent to which innovations and thereby knowledge diffusion are covered by patent data varies significantly between industries (Arundel and Kabla, 1998). This makes a comparison of patent-based observations difficult across industries. In order to minimize the likelihood of an inter-industry bias in the use of patent data as indication of knowledge diffusion, we decided to consider exclusively technology specific patent citations excluding citations between technologies. Future studies might be able to tackle this issue in other ways and, hence, might be able to exploit the full breath of patent citations. In addition, future research should make use of alternative data sources to measure knowledge diffusion. One way could be to relate firms’ product portfolio diversification to their cooperation activities, as their diversification might be the consequence of obtaining and transforming new knowledge. The chapters’s fourth shortcoming is the consideration of only one specific type of subsidy, i.e., that granted by two ministries of the German federal government. Hence, our findings remain restricted to these programs, as other subsidization schemes such as the EU-Framework Programmes or local policy initiatives might induce different processes and be more effective. Clearly, more research is needed in this direction.

In addition to these shortcomings, there might be other processes that might be responsible for our finding of R&D subsidies not enhancing inter-regional knowledge diffusion. It has been frequently shown that joint R&D projects tend to bring together organizations in geographic, cognitive, organizational, social, and institutional proximity (Breschi and Cusmano, 2004; Scherngell and Barber, 2009, 2011; Balland, 2012; Broekel and Hartog, 2013a, 2013b). Yet, these are precisely those constellations that are most likely to emerge without subsidization. It can therefore be argued that to stimulate knowledge diffusion, policy should try to stimulate interactions among organizations that are rather unlikely to interact in the first place. For instance, if subsidized collaboration primarily connect partners at low cognitive distances, their learning potentials are rather small, implying relatively low possibilities for mutual referencing (e.g. in form of patent citations) because they are already familiar with each other's work. Similar arguments apply in the case of organizations sharing a long history of collaboration, which also seem to team up frequently in subsidized joint R&D projects (Breschi and Cusmano, 2004).

In summary, this chapter marks an additional step towards a better understanding of the effects and effectiveness of subsidizing joint R&D projects. Yet, as the discussion shows, the study opens a barrage of additional issues that hopefully will be addressed by future research.

Appendix

3.A1 Additional analysis (five-year time lag)

Table A3.1: Results of the fixed-effects, binomial and negative binomial (count) regression models (5-year time lag)

Variables	Binomial model 6 Estimate (SE)	Count model 6 Estimate (SE)
Policy		
$SUBS.NET_{i,j,(t-5),s}$	0.024	-0.009
	(0.018)	(0.009)
$SUBS_{i,t,s}$	0.013***	0.005***
	(0.003)	(0.002)
$SUBS_{j,t,s}$	0.02***	0.002
	(0.003)	(0.002)
Proximities		
$CO\_INV_{i,j,(t-5),s}$	0.11***	0.012***
	(0.005)	(0.001)
$DIST_{i,j}$	-0.098***	-0.017***
	(0.004)	(0.003)
$TECH_{i,j}$	1.809***	0.124***
	(0.038)	(0.027)
$ORG_{i,j}$	-0.002***	0.006***
	(0.004)	(0.001)
Regional properties		
$PAT_{i,t,s}$	0.018***	0.004***
	(0.0005)	(0.0003)
$PAT_{j,t,s}$	0.015***	0.003***
	(0.0005)	(0.0003)
Control		
$CIT\_lag_{i,j,t}$	-0.236***	0.047***
	(0.011)	(0.003)
Year Dummies	Yes	Yes
Technology Dummies	Yes	Yes
AIC	163808	102700
2x Log-likelihood		-102532
Observations	255493	34487
Non-zero obs.	34487	-

Disentangling link formation and dissolution in spatial networks

Abstract

The analysis of spatial networks' evolution has predominantly concentrated on the formation process of links. However, the evolution of networks is similarly shaped by the dissolution of links, which has thus far received considerably less attention. The chapter presents separable temporal exponential random graph models (STERGMs) as a promising method in this context, which allows for the disentangling of both processes. Moreover, the applicability of the method to two-mode network data is demonstrated.

We illustrate the use of these models for the R&D collaboration network of the German biotechnology industry as well as for testing for the relevance of different forms of proximities for its evolution. The results reveal proximities varying in their relative importance for link formation and link dissolution.

This chapter is co-authored with Tom Broekel. The PhD candidate is the second author of the article. The paper has been published as "Disentangling link formation and dissolution in spatial networks: An Application of a Two-Mode STERGM to a Project-Based R&D Network in the German Biotechnology Industry" in Networks and Spatial Economics, 2018, 18, 677–704 https://doi.org/10.1007/s11067-018-9430-1

4.1 Introduction

Network analysis has gained great popularity in many spatial disciplines (Ducruet and Beauguitte, 2014). For instance, in urban studies, network analyses are intensively used to study city-networks (Liu et al., 2013), while economic geography focuses on R&D networks' facilitating of the flow of knowledge between cities and regions (e.g., Murphy, 2003; Boschma and ter Wal, 2007). In both fields, studies have sought to explain the evolution of inter-organizational relationships in time and space by relying on longitudinal network data (Broekel et al., 2014). Most of the existing research focuses on the relative importance of factors facilitating link formation. Crucially, network evolution consists of link formation and dissolution processes, though different factors might drive each process. For instance, Balland (2012) note "[…] that the creation and dissolution of ties are not generally strictly inverse mechanisms […]" (p. 749). Moreover, Krivitsky and Handcock (2014) explain that "social processes and factors that result in ties being formed are not the same as those that result in ties being dissolved" (p. 35). For instance, in order to benefit from scale effects, firms might participate in joint R&D projects with other firms that have a similar technological background (i.e., they are cognitively proximate). Over the course of the project, they realize that their technological similarity stimulates unintended knowledge spillovers, and they end the collaboration to sustain their competitive advantages. Hence, cognitive proximity fostered collaboration in the first place and subsequently increased the likelihood of an early termination of the collaboration. However, while substantial empirical evidence of the first process exists, much less attention has been paid to the second process.

The present chapter contributes to the spatial network literature in two ways. Firstly, it demonstrates the use of separable temporal exponential random graph models (STERGMs) as a method for investigating formation and dissolution processes in spatial (knowledge) networks (Krivitsky and Handcock, 2014). We apply STERGM to a spatial network emerging from subsidized R&D projects in the German biotechnology industry between the years 1998 and 2013. Secondly, we demonstrate STERGM's ability to handle two-mode network data, which overcomes the (still) common but sometimes questionable one-mode project of network data when constructing spatial (knowledge) networks (Scherngell and Barber, 2009, 2011; Balland, 2012; Broekel and Hartog, 2013b; Hoekman et al., 2013; Buchmann and Pyka, 2015). We thereby extend the work of Liu et al. (2015), who applied a cross-sectional two-mode exponential random graph model to analyze global city networks by presenting an application of ERGMs to longitudinal data. While, alternatively, such data can be investigated with stochastic actor-oriented models (SAOMs) (Liu et al., 2013), these models require specific assumptions (e.g., agency) that are often doubtful in the context of spatial networks (Broekel et al., 2014). In addition, STERGMs have been shown to be empirically similar if not preferable to SAOM models (Leifeld and Cranmer, 2015).

This chapter is organized as follows: Section 4.2 discusses the process of an inter-organizational R&D cooperation network evolution. It addresses the relevance of organizations' attributes, their relational characteristics, and structural level effects. It also considers why existing empirical analyses on their relative importance might be biased, which motivates the use of STERGMs. The STERGM approach is introduced in Section 4.3. Section 4.4 discusses the network data and the empirical model specification. The analyses' results are presented and discussed in Section 4.5. Section 4.6 concludes the chapter.

4.2 Disentangling the determinants of link formation and dissolution

On the following pages, we will argue why we expect the influence of factors to vary for formation and dissolution processes, whereby varying effects are particularly likely for proximities and the location.

The literature on the evolution of spatial networks generally highlights three essential levels at which processes of network evolution occur (Glückler, 2007; Ter Wal and Boschma, 2009; Boschma and Frenken, 2010). These levels are the (1) the node, (2) the dyad, and (3) the structural network.

4.2.1 The node level

Many organizational characteristics influence the collaboration behavior of organizations. Researchers have often argued that the size of organizations is of relevance: In particular, two of the expressed arguments are in favor of greater nodes having more links. First, larger nodes, i.e. organizations with more employees, may have greater capacities to establish and maintain more links (Tether, 2002). Second, larger nodes tend to attract more requests for interacting, as they usually occupy more prominent positions within specific fields in general and within existing networks (Broekel and Hartog, 2013a). For instance, larger organizations are more widely known than smaller ones and, due to their larger portfolio, provide more opportunities for interacting. In case of the Dutch aviation knowledge network, Broekel and Hartog (2013b) find evidence for a positive relationship between size and link formation.

Moreover, larger firms might have more capacities to form new relationships and simultaneously maintain previously existing ones. Smaller firms tend to face a trade-off in this situation — i.e., they must decide whether to invest time in establishing new relationships and giving up existing ones or to opt for maintaining their relations (Tether, 2002). Hence, small organizational size might negatively relate to link formation and dissolution, while in case of large organizations, the latter relationship might be positive. This point also highlights that link formation and dissolution are not necessarily independent of each other because of organizations' potential constraints in their collaboration capacity.

In the literature on spatial (knowledge) networks, organization-specific characteristics (nodes) are complemented by factors at the spatial level, which also impact organizations' interaction behavior. For instance, Illenberger et al. (2013) hypothesize differences in the relationship structures of individuals living in cities and those in rural areas. While they failed to empirically confirm this hypothesis, empirical evidence exists for organizations in urban and rural areas. For example, Meyer-Krahmer (1985) report that firms in (urban) agglomerations are more prone to interact with other organizations than firms in rural areas. Broekel and Hartog (2013a) confirm this positive relationship between population density and organizations' amount of inter-regional collaboration. Moreover, Wanzenböck et al., (2015) investigate the centrality of regions in inter-organizational R&D networks initiated by the EU Framework Program. Their findings clearly show urban regions being more central in these networks than rural regions. Hence, as an example of a spatial factor influencing organizations' interaction behavior, we focus on organizations' location within urban regions, which is expected to facilitate their link formation activities.

In regard to link dissolution, we further argue that these positive urbanization externalities (Boschma and Wenting, 2007) will help organizations to maintain relationships. By accessing major train stations and airports, organizations tend to be able to lower transportations costs and will be able to maintain more relationships than organizations situated in more remote rural areas.

> **Hypothesis 1**: Organizations located in urban areas are more likely to form a link and less likely to dissolve a link.

4.2.2 The dyad level: How proximities shape network structures

The dyad level refers to the properties of the relationships between nodes. In research on spatial networks, Boschma's (2005) proximity framework offers an effective summary of many (specific) arguments made in the literature. Among others, the concept builds upon the homophily effect, which has been applied in sociology. Here, it is argued that two individuals are more likely to develop a trust-based relationship when they share similar attributes (e.g., the same age) (McPherson et al., 2001). This concept has been transferred to the organizational and regional levels as well as to other types of relationships and similarities. More precisely, Boschma (2005) summarizes the prominent arguments in the literature and proposed a distinction between five dimensions of inter-organizational proximities. These proximities describe organizations' similarity (homophily) in different dimensions and are all argued to increase the likelihood of two organizations to establish a (collaborative) relationship and to exchange knowledge. These proximities are cognitive, geographical, organizational, social, and institutional.[8] As our empirical analysis will focus on cognitive, institutional, and geographic proximity, we limit the theoretical discussion to these dimensions. A discussion on the other two dimensions can be found in Boschma (2005).

Nooteboom et al. (2007) define cognitive proximity as the result of organizations' development of an organization-specific internal "interpretation system" (p. 1017). At its core is the organizations' absorptive capacity. As learning is a cumulative process that builds upon existing knowledge, their absorptive capacity increases when new and previously possessed knowledge overlap (Cohen and Levinthal, 1990). Accordingly, organizations tend to interact with partners who share similar knowledge bases. In this case, it is easier and more efficient for organizations to identify them as potential collaboration partners, absorb their knowledge, and jointly learn (Nooteboom et al., 2007). The positive impact of cognitive proximity on link formation in spatial R&D networks has been frequently confirmed (Paier and Scherngell, 2011; Balland, 2012; Broekel and Hartog, 2013b; Buchmann and Pyka, 2015).

While cognitive proximity greatly increases link formation, two cognitively similar organizations are likely to be competitors because they tend to produce similar products (Boschma, 2005). This circumstance increases the risk of withholding knowledge in order to avoid unintended knowledge spillover (Zander and Kogut, 1995). Moreover, given their cognitive overlap, these organizations offer relatively little to learn from each other. In such a situation, the formed alliance may be unstable (Polidoro et al., 2011), as organizations tend to be reluctant to stay in alliances longer than necessary. Accordingly, cognitive proximity may increase the chances of early link dissolution.

Geographical proximity refers to the "similarity" of organizations in terms of their geographic location. Being geographically close or within the same region fosters the formation of links because it makes frequent face-to-face interactions much easier (Boschma, 2005). Such contacts facilitate the generation of mutual trust and are especially important when exchanging

[8] This list of proximities is not exclusive. Other types of proximities may matter as well but have received considerably less attention in the literature so far.

tacit knowledge (Ter Wal, 2014). In spatial sciences, geographical proximity is a key interest and thus is often analyzed in regard to network formation. For instance, in the case of funded R&D networks, Paier and Scherngell (2011) as well as Balland (2012), among others, find evidence of a positive relationship between link formation and geographical proximity.

As geographic proximity strongly enhances the possibility of frequent face-to-face contacts and more insightful communication, it may contribute to the earlier completion of projects, which in turn will result in quicker link dissolution. It might even be the case that partners anticipate the higher efficiency and more effective communication when collaborating in geographic proximity and therefore opt for shorter project durations when setting-up collaborations with geographically proximate partners, such as when applying for joint grants.

Hypothesis 2: Geographic and cognitive proximity positively influences link formation and dissolution.

Institutional proximity is also associated with the embeddedness literature, i.e. organizations operating in different social subsystems (e.g., industry or academia). According to Ponds et al., (2007), scientific research and the development of product innovations are "conducted within different socio-economic structures" (p. 426). Institutionally distant organizations are more likely being confronted with unknown behavior and problems in mutual communication, which reduces the likelihood of interaction (Parkhe, 1991; Boschma, 2005; Balland et al., 2013). Institutional proximity ensures that partners operate under the same or at least comparable institutional (legal and societal) frameworks, which significantly aids in overcoming the risks of freeriding and reduces monitoring costs (Boschma, 2005). Accordingly, it strongly helps with initiating collaborations, which is also empirically confirmed (Balland, 2012).

In contrast, its relevance for link duration might be rather minimal. It can be argued that once collaboration has been initiated and formalized, most legal and formal issues concerning the collaboration are settled and contractually fixed. While the efforts needed for this may prevent the formation of interactions, the institutional frameworks may become complementary through the formal contract and, hence, exercise little to no effect on link duration.

Hypothesis 3: Institutional proximity impacts link formation positively but does not affect link dissolution.

4.2.3 Structural level determinants

Glückler (2007) and Liu et al. (2015) highlight the relevance of factors at the structural network level. These authors argued that a theory of network evolution focuses on the interdependency of new links and the overarching structure of the network as such. Accordingly, "[…] this perspective explicitly moves beyond the dyadic analysis of single relationships to the analysis of entire network relations" (Glückler, 2007: p. 207). Three factors have received the most attention so far: triadic closure, multi-connectivity, and preferential attachment (Ibid.).

Triadic closure implies that partners of a node are likely to become partners themselves. This is shown by so-called triangles in networks, i.e. dense cliques of strongly interconnected nodes (Ter Wal, 2014). In spatial (knowledge) networks, such cliques are usually interpreted as a sign of social capital (Coleman, 1988), which may enhance trust and the willingness among nodes to invest in mutual goals. For instance, Ter Wal (2014) confirms the relevance of triadic closure for the evolution of a biotech network based on co-invented patents.

Multi-connectivity is a consequence of organizations tending to seek a diverse portfolio of partners. In other words, they may connect to others in multiple ways to decrease their dependency on individual links (Glückler, 2007). For example, organizations may link to other organizations through joint R&D projects in addition to existing buyer-supplier relations. Broekel and Hartog (2013b) provide empirical evidence for the relevance of such processes in the context of subsidized spatial networks.

Preferential attachment implies that the probability of creating additional links may increase with every new link a node possesses (Vinciguerra et al., 2010; Liu et al., 2015). Organizations with many relationships tend to have a greater flow of information about new activities and partners, and they also tend to have a stronger ability to evaluate these by means of collaborative behavior and appropriate resources (Polidoro et al., 2011). While Broekel and Hartog (2013a) hypothesized preferential attachment to play a role in networks of subsidized R&D collaboration, they failed to empirically confirm this. With respect to their relevance on link formation and dissolution, the literature clearly suggests a positive contribution to link formation, while discussions on their effects for link dissolution are largely absent. We therefore expect that their positive influence is also applicable to link persistence (i.e. these effects are negatively correlated with link dissolution).

Hypothesis 4: Network structures support link formation and suppress link dissolution.

4.3 Separable temporal exponential random graph models

A range of methods can be applied to identify factors driving networks' evolution (see, for example, a recent review of the most common approaches: (Broekel et al., 2014). In the context of dynamic spatial networks, SAOM models in particular have been used (Balland, 2012; Liu et al., 2013). These models are convincing due to their wide range of application possibilities, consideration of factors at all three levels of investigation, and usability with one and two-mode network data. While their applicability and functionality were unmatched in the past, the development of the TERGM (temporal exponential random graph model) and STERGM (separable temporal exponential random graph model) provides researchers with a legitimate modeling alternative. It is beyond the scope of the present chapter to conduct a full review and an empirical comparison of the two models. For this, we refer to Broekel et al., (2014) and even more so to Leifeld and Cranmer (2019). The chapter instead focuses on an application of the recently developed STERGM and seeks to highlight its three most prolific features that are crucial in the context of spatial (knowledge) networks: its nature as a tool of dynamic network analysis, its applicability to two-mode network data, and its ability to separate formation and dissolution processes. While SOAMs offer similar features, these are achieved by the fundamental assumption of agency residing with the nodes. In other words, the models are built on actor-based behavioral assumptions (Park and Newman, 2004). When applying these models to inter-organizational or inter-regional networks, this assumption of agency is likely to be violated (Broekel et al., 2014). Moreover, recent theoretical and empirical comparisons suggest that (S)TERGMs outperform SOAMs (Leifeld and Cranmer, 2015), which further motivates the presentation of STERGM for the analysis of spatial (knowledge) networks.

The separable temporal exponential random graph model (STERGM) is a recently developed extension of the exponential random graph model (ERGM) (Krivitsky and Handcock, 2014); as such, it is part of the ERG family (also known as p*-models (Robins et al., 2007)).

As neither nodes (actors) nor dyads (relationships) are completely independent from each other, classical econometric models such as regression analysis do not effectively explain the structure of observed networks (Broekel et al., 2014). For that reason, Frank and Strauss (1986) develop the so-called Markov dependence on which ERGMs are based. It implies that a given dyad between two actors impacts and is impacted by any further link of those two actors (Robins et al., 2007). Therefore, links are defined as being "conditionally dependent" (Ibid: p. 181).

Models of the ERG family consider link creation as a continuous process, and the observed network structure is seen as one possibility out of a large set of potential networks with similar characteristics (Ibid.). This range of possible network patterns and their likelihood of appearance "is represented by a probability distribution on the set of all possible graphs with this number of nodes" (Ibid: p. 176). Hence, a good ERGM has a high probability of simulating the observed network by finding the correct coefficients of the determinants impacting the network structure. For this purpose, a Markov chain Monte Carlo maximum likelihood estimation (MCMC MLE) procedure is used to simulate and evaluate the modeling process (Broekel et al., 2014).

Mathematically, an ERGM is defined as follows (Robins et al., 2007):

$$Pr(Y = y) = \left(\frac{1}{\kappa}\right) exp\{\textstyle\sum_A \eta_A g_A(y)\}$$

where $Pr(Y = y)$ is the probability that the observed network (y) equals the simulated network (Y). The network configuration A is considered by η_A, and the network statistics are represented by $g_A(y)$. The network configurations are the determinants with which the researcher attempts to explain the network structure, such as cognitive proximity. ERGMs allow the inclusion of node, dyad, and structural determinants at the same time (Broekel et al., 2014). $g_A(y)$ is either 1 if the configuration is observed in y, or 0 if it is not. The factor κ is a normalizing constant that is implemented to ensure a proper probability distribution of the equation (Robins et al., 2007).

Hanneke and Xing (2007) and Hanneke et al. (2010) extend the ERG family with a framework that enables the researcher to model network dynamics over discrete time steps, called temporal ERGM (TERGM). In this model, a network at time t is conditional on the network at time t - 1. In essence, the TERGM corresponds to a stepwise ERGM approach with the steps corresponding to the observed time periods (Krivitsky and Handcock, 2014). Recently, Krivitsky and Handcock (2014) build upon this model and introduced the concept of separability. This allows a STERGM to independently consider the process of link formation and dissolution. In consideration of the organizational processes underlying the establishment and maintenance of cooperation, it seems legitimate to view different factors as in control of link formation and dissolution. A STERGM displays the transition from one time period (t) to the following time period (t+1) and thereby independently analyses the formation and dissolution of links. Accordingly, a STERGM is separated into two formulas (Ibid.). One formula considers the formation of links:

$$Pr(Y^+ = y^+|Y^t) = \left(\frac{1}{\kappa^+}\right) exp\,\{(\eta_A{}^+)^t g_A(y^+)\}$$

The other formula considers the dissolution of links:

$$Pr(Y^- = y^- | Y^t) = \left(\frac{1}{\kappa^-}\right) exp\,\{(\eta_A{}^-)^t g_A(y^-)\}$$

The general aim of this method is to obtain a model with a high probability of simulating the observed network and that can identify the best coefficients. The success of the simulation can be tested by checking whether the model is degenerated and by examining the model's goodness of fit. A degenerated model is often the consequence of misleading starting parameters and/or variables that are not able to correctly simulate the observed network. A degenerate model does not converge, or the calculated estimates simulate a network that is either extremely dense or has almost no edges (Robins et al., 2007).

A non-degenerated model has to be further tested regarding the quality of simulating the observed network. By comparing the network characteristics of the simulated network (e.g., the degree distribution) with the corresponding statistics of the observed network, the goodness of fit can be verified graphically (Hunter et al., 2008).

When calculating several models of the same size but with slightly different variables, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) provide additional information on a model's goodness of fit. However, by including several network configurations (variables), the model becomes increasingly complex, and both AIC and BIC become less precise (Goodreau, 2007). Therefore, they should only be used in combination with the graphics mentioned above.

We take advantage of the STERGM being capable of handling two-mode data. Accordingly, a one-mode projection is not necessary, but we directly analyze the two-mode structure of the network. In this case, the researcher must make sure that the simulation procedure does not create links that are impossible, i.e. no links should be simulated among events or among participants, only between events and participants (for practical application see Morris et al. (2008) and Section 4.4.4).

4.4 Empirical approach and data

4.4.1 Data

The empirical network is based on organizations' participation in joint R&D projects subsidized by the German Federal Ministry of Education and Research (BMBF), the Federal Ministry of Economics and Technology (BMWi), and the Federal Ministry of the Environment, Nature Conservation and Nuclear Safety (BMU). Data on subsidized R&D projects are extracted from the so-called “Förderkatalog” (subsidies catalogue)[9]. Financial support for joint R&D projects is conditional on all participants agreeing to exchange knowledge with each other. Moreover, they grant access to intellectually property rights that are within the scope of the project but existed before project’s start (BMBF, 2008). Therefore, inter-organizational relations based on joint participation in such subsidized projects qualify as knowledge exchange links (Broekel and Graf, 2012). The data consist of firms, universities, and research institutes that operate in the German biotechnology industry and obtain subsidies for their joint projects in the period from 1998 to 2013.

The industry has been chosen because it can be classified as a science-based industry in which scientific advancements primarily drive economic progress (Ter Wal, 2014). Moreover, cooperation is essential for innovation in this industry, as its “locus of innovation” is located in the network of inter-organizational relationships rather than in a single organization (Powell et al., 1996: p. 119). Thus, inter-organizational R&D cooperation is an important competitive factor in the biotechnology industry because individual firms may not be able to cover all the necessary capabilities to innovate (Ibid).

Regarding the economic entities being used as nodes in the network analysis, the subsidies catalogue distinguishes between the beneficiary unit (“Zuwendungsempfänger“) and the executing unit (“Ausführende Stelle“). The first refers to the receiving organization (e.g., organizations’ headquarters), and the latter refers to the executing entity (e.g., a specific department or an institute of this organization). In accordance to the literature (Broekel and Graf, 2012), we chose the executing units as network nodes because they actively select whom to cooperate with and decide when to end a project.

4.4.2 The structure of two-mode networks

The described data represent a two-mode (or bipartite) network, as actors are related to projects and not directly to other actors. We extracted 652 nodes at the actor level (mode 1; i.e., organizations) and 258 nodes at the event level (mode 2; i.e., projects). Both levels are connected through 1,177 links (see Figure 4.1). The two-mode network structures have significant implications for network analysis, as, for instance, network structures such as closed triads are not possible.

[9] In addition to the subsidies catalogue, the websites “Biotechnologie.de,” “chemie.de,” “Life-Sciences-Germany.com,” and “statista.de” and the homepages of the organizations have been used to acquire further data on organizational size and technological focus (cognitive proximity).

To account for factors' importance varying over time (see, e.g., Balland et al., 2013), we split the network into four phases, with each being four years (see Figure 4.1 and Table 4.1). We defined a link to be formed when a project started within the observed time phase. It was maintained when the project had not been ended during the foregoing timespan. Otherwise, the link was been dissolved (see Figure 4.2). Pooling the data for four years caused the resulting networks to be sufficiently dense. We analyzed the three transitions of the networks from one period to the next by estimating separate models for each transition. This allowed for assessing potentially time-varying effects of our explanatory variables.

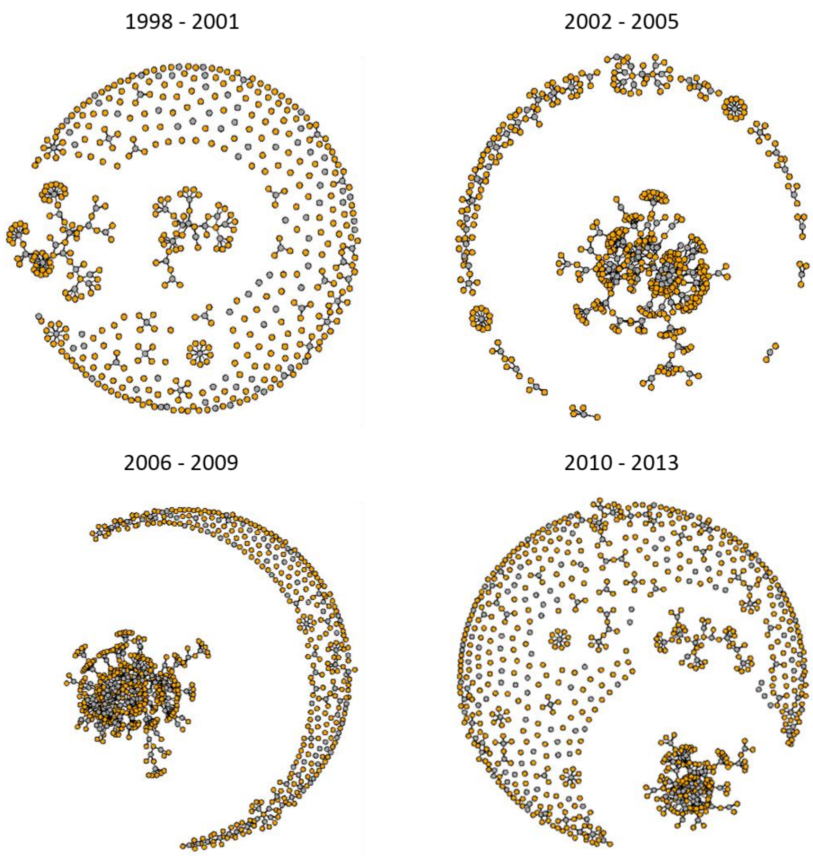

Figure 4.1: Network visualization for all four time periods (grey = projects, orange = organizations)

The STERGM demands the network to have the same set of nodes in both time periods. This gave us two opportunities: First, we could have included all nodes in the networks, regardless of whether they have a link in that period. However, this would have led to more complex models and would have decreased the chances of a converging model. Moreover, nodes only participating in the first transition are irrelevant for the following transitions. Therefore, we went with the second possibility: In the first STERGM, we only considered nodes that participate in the first and second periods. In the second STERGM, we then only included nodes participating in the second and third periods. Finally, in the third model, we only considered nodes that had a link in the third or fourth period. Eventually, we had two slightly different networks for the second period and the third period (see Table A4.1 in the appendix for an overview of the networks).

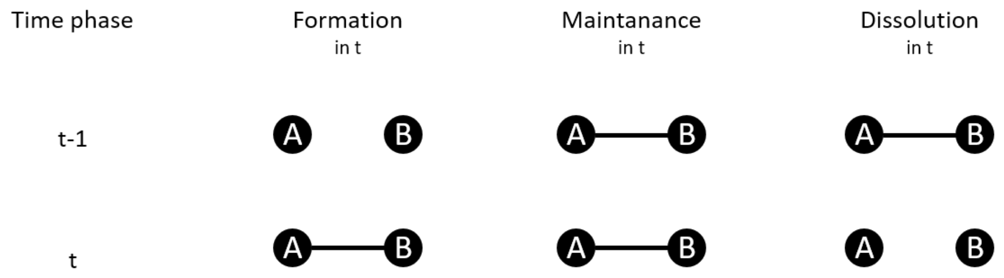

Figure 4.2: Link formation, maintenance and dissolution (a = organization, b = project)

Table 4.1: Link development from 1998 – 2013

Network	Period	Links present	Links formed	Links maintained	Links dissolved
No. 1	1998 – 2001	236	236	-	-
No. 2	2002 – 2005	494	275	219	17
No. 3	2006 – 2009	689	465	224	270
No. 4	2010 – 2013	530	145	385	304

In the case of publicly subsidized project data, multiple reasons may exist for the dissolution of links. First, if participants successfully complete the project within the subsidized time period, the network link(s) will disappear. Second, if organizations apply for and receive a second funding within the project run-time, the link will be extended without a break, and we would not observe the dissolution of a link. Interestingly, we did not find a single instance in which this took place. We speculated that a policy discriminates against immediately reoccurring project partnerships when awarding new grants. Third, a policy could artificially induce the termination of joint projects and the according dissolution of network links, thus setting a maximum project duration. While this motivates Balland (2012) to argue that "analyzing why links are dissolved […] in the case of projects whose length is fixed from the beginning seems less relevant" (p. 749), we argue that partners know about fixed project durations ex-ante. Hence, they will apply for a grant only if its duration meets the (foreseeable) requirements of the planned project, which includes the consideration of the scope, complexity, and partner characteristics. Each of these considerations is usually known ex-ante to some extent. Similar to Makino et al. (2007), we therefore expect the initial conditions of partner selection to influence the projects' length. For instance, we expect more complex (and therefore longer) projects to more likely involve geographically proximate partners, as the complexity requires more frequent face-to-face contacts (see, e.g., Balland and Rigby (2017). Similarities can be expected for projects involving actors at greater cognitive distances, which also tend to demand increased and closer interaction (Boschma, 2005). Two processes are likely to support this. Firstly, when designing subsidization programs, a policy is probable to consider the task's complexity and defines longer project durations. Secondly, applicants may look for programs with maximal project durations that fit the complexity of the expected task. We assumed project-lengths are (indirectly) related to the type of partners and consortia applying. Significant results in the dissolution models will show the extent to which this assumption is valid.

Based on these arguments and secondary data, we constructed the following variables at the node, dyad, and structural network levels.

4.4.3 Dyad level variables

Categorical and binary dyad-level effects are considered in the STERGM by evaluating how frequently two-paths are created between two organizations sharing the same characteristics (see Figure 4.3). We were thereby particularly interested in their characteristics concerning cognitive, geographical, and institutional proximity. We did not consider social and organizational proximity because of missing data.[10]

In the biotechnology industry, organizations are commonly assigned to a technological subfield: medicine and pharmacy, industrial processes, agriculture, and (bio)informatics (DaSilva, 2012). These fields represent distinct technological foci and systematic differences in the way R&D is conducted (Herrmann et al., 2012). We constructed a simple measure of cognitive proximity based on this assignment. If two partners were assigned to the same category, they were perceived of as being cognitively more proximate than in the case they were active in different technological subfields. The variable *COG PROX* was given a value from 1 to 4 according to the assigned subfields[11].

[10] In general, the data allowed us to compute organizational proximity because of the distinction between beneficiary and executing entity. If two collaborating entities were departments of the same beneficiary, they would have a higher organizational proximity. However, in the data set at hand, this setting is extremely rare (around 1%) and, thus, very likely to be insignificant anyway.

[11] Unfortunately, we could not assign a biotech subfield to every organization (see Table A4.2 in the appendix). Fortunately, the STERGM allows for excluding categories from the calculation, which we made use of when calculating the effect of cognitive proximity.

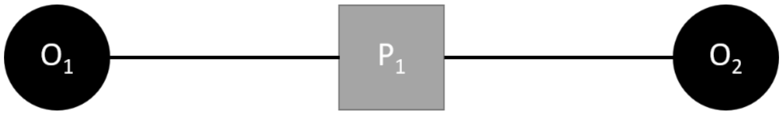

Figure 4.3: Homophilous two-path of organization 1 and 2 via Project 1

The measure of geographic proximity (*GEO PROX*) is a categorical variable corresponding to the NUTS 3 region in which organizations are co-located. In Germany, NUTS3 regions correspond to 429 districts (Kreise), which are administrative areas ranging from cities such as Munich or Berlin to rural areas such as the Uckermark in East Germany (for additional Figures, see Table 4.2).

Table 4.2: Figures of German districts (NUTS 3) (Source: Destatis, 2016)

	Average	**Minimum**	**Maximum**
Area (km²)	907	35	5,470
Population	203,589	34,260	3,520,031
Population density (per km²)	504	36	4,668

Moreover, organizations were classified as being profit orientated (private firms) and as non-profit organizations (universities, research institutes, and associations). This difference was captured by our measure of institutional proximity (*INST PROX*), which is categorical and distinguishes between firms (0), universities (1), and research institutes (2).

4.4.4 Organizational node level variables

Potential location effects of organizations situated in urban areas were approximated using data of the Federal Institute for Research on Building, Urban Affairs and Spatial Development. It classifies each German NUTS 3 regions as “urban,” “increasing urbanization,” or “rural.” The classification is based on the total population and population density (BBSR, 2015). We constructed the categorical variable (*URBAN*) as 0 for rural, 1 for increased urbanization, or 2 for urban regions.

The second variable at the node level approximated the size of organizations. As it was impossible to acquire the number of employees for each organization and year, we created a categorical variable (*SIZE*) indicating membership in different size classes. SIZE consisted of the categories utilized by the Reconstruction Credit Institute (KFW, 2012) as well as Buchmann and Pyka (2015):

Category 0: organizations with fewer than 50 employees.

Category 1: organizations with 51 to 250 employees.

Category 2: organizations with more than 250 employees.

The third node level variable is *EAST*, which distinguishes organizations located in West (Category: 0) and East Germany (Category: 1). To the catching-up process of the East German economy, a large share of European and German subsidies is allocated there to facilitate this process. Thus, there might be a propensity to favor applications from organizations being located in cities formerly belonging to the German Democratic Republic (GDR). Moreover, Cantner and Meder (2008) discover that East German organizations participate more actively in R&D collaborations.

As we sought to model interactions between specific variables (see Section 4.3.), we also considered the corresponding main effects at the node level. We therefore included node-level variables consisting of the categories of cognitive proximity (i.e., *MEDICINE* and *AGRICULTURE* with base *INDUSTRIAL*[12]) and the differentiation between types of organizations (i.e., *UNI* and *RESEARCH INST* with base *FIRM*)[13]. While surely being interesting on their own, due to the scope of the study, we primarily included these variables as control variables.

4.4.5 Structural level variables

At the structural level, four variables were considered.[14] The effect of multi-connectivity was captured by the so-called geometrically weighted dyad shared partner statistic (*GWDSP*). A positive coefficient of this statistic suggests that actors tend to link in multiple ways (i.e., via multiple projects) to each other (Hunter et al., 2008).

The second structural determinant is preferential attachment. We modeled this by making use of the variable *GWDEGREE*, which represents the geometrically weighted degree statistic. The variable is seen "as a sort of anti-preferential attachment model term" (Hunter, 2007: p7).

[12] Bioinformatics was excluded as only 25 organizations are assigned to this category over the complete timespan.

[13] As the categories of GEO PROX consist of approximately 80 regions, we excluded them as well, as it would have made the models too complex to calculate.

[14] Our two-mode network has no triads and STERGM currently does not support the consideration of a two-mode clustering coefficient as, e.g., described by (Opsahl, 2013). We will therefore not further elaborate on triadic closure, which does not mean that it is of no relevance.

If its coefficient is negatively significant at the actor level[15], preferential attachment is a likely driver of network evolution. In contrast, there is no clear interpretation of a significant coefficient of *GWDEGREE* at the event level. It means that preferential attachment works at the project level, which lacks a theoretical foundation. Nevertheless, the effect was included to help the simulating of the network.

The observed networks are characterized by high numbers of projects with three participants (see Figure 4.5). We considered this by including the variable *B2DEG3*, which added a statistic to the model counting how frequently B2-nodes (projects) have three links, i.e. three participants (Morris et al., 2008).

The final structural network variable is *EDGES*. This variable should always be included when modeling a network with any ERG method. It equals the number of observed edges and helps in modeling the density of the observed network in the simulations (Broekel and Hartog, 2013b).

In the appendix, Table A4.2 presents the descriptives of all node and dyad level variables.

4.5 Results and Discussion

4.5.1 Verifying the model

Before presenting the empirical results, it is important to address a number of issues that have to be taken into consideration before interpreting the results. For instance, there might be a potential bias connected to our data. For historic reasons, subsidized R&D projects frequently (but not exclusively) have a length of 36 months (see Figure 4.4).

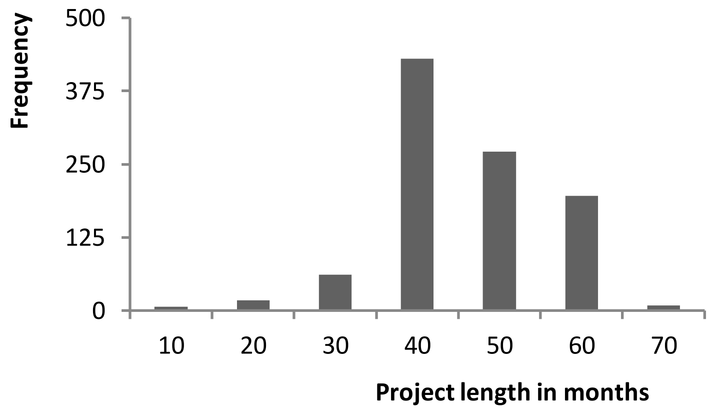

Figure 4.4: Frequencies of project length in months, n = 750

[15] STERGM allows for calculating GWDEGREE for both levels (actor and event). A significantly negative coefficient will be obtained if the network shows a power law degree distribution. It means that at the actor mode, few organizations participate in many projects. At the event mode, few projects have many participants in this case.

Accordingly, project lengths are not fully flexible, and organizations do not have full freedom in choosing a support scheme allowing for project lengths that meet their requirements (also see Section 4.4.2). In other words, this precondition dominates link dissolution. To evaluate the significance of this, we created a second network that eliminated all links of projects that terminated immediately after 36 months. Projects and participants that became isolates because of this circumstance were also deleted. The corresponding network consisted of 144 projects and 476 actors.

There are two implications. Firstly, due to the predefined project lengths, we were less likely to obtain significant coefficients in the dissolution model, as project endogenous processes and conditions are “overruled” by these externally imposed conditions. In other words, link dissolution becomes an external event and hence cannot be explained by endogenous processes. Secondly, if significant coefficients are obtained or differences between the models for the full set of projects and those excluding links of 36 months are observed, these should be primarily interpreted as selection effects — i.e., partners choose specific support schemes considering the maximal time of subsidization when applying for grants.

In general, the results do not change significantly when excluding the 36-month projects, which indicates, similar processes drive both networks’ evolution. A major difference is related to geographic proximity. It was not possible to find a converging model when considering the full set of projects. However, when excluding the 36-month projects, convergence was achieved, and we obtained reliable results.

Besides convergence, STERGM involves finding the best model in a manual iterative trial-and-error process (Broekel and Hartog, 2013b). Usually, a first estimation is used to calculate starting values entering the second estimation (similar to Goodreau (2007). The models’ goodness of fit is assessed via the degree distribution. Figures 4.5 and 4.6 plot the observed network’s degree distribution as a solid line and the 95% confidence interval of the distribution for the corresponding simulated networks as box-plots and light-grey lines. A solid line within the light-grey lines represents a model with a satisfying goodness of fit (Krivitsky and Goodreau, 2015). The figures reveal our models as being of sufficient overall quality because only small parts of the simulated degree distribution exist outside of the observed one (Ibid).

The coefficients of the formation and dissolution model can be understood as odd ratios by taking the exponential. In the case of the formation model, a positive coefficient means that the establishment of a link is more likely. In contrast, in the dissolution model, a positive sign signals persistence of a link, i.e. the lower likelihood of dissolving (c.f. Krivitsky and Goodreau, 2015).

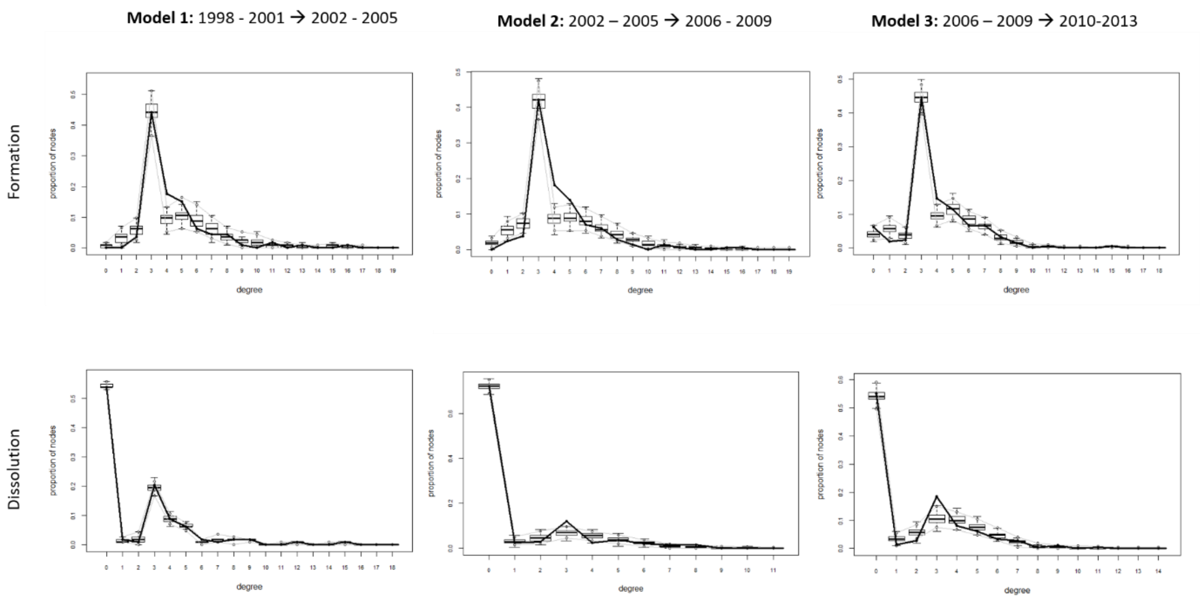

Figure 4.5: Degree distribution of all the initial models

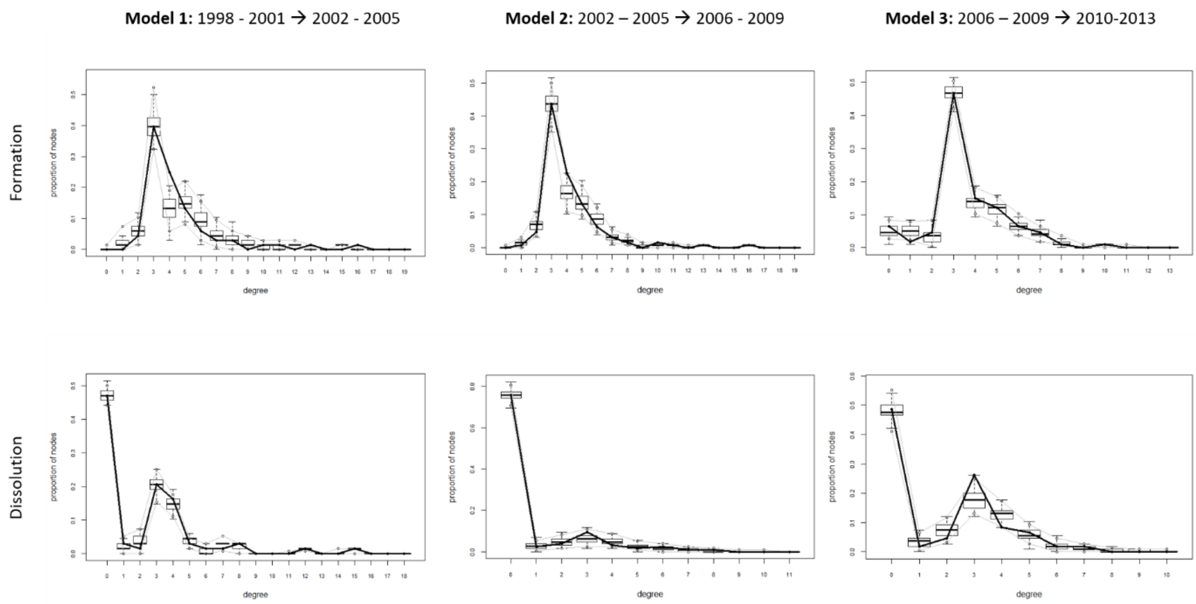

Figure 4.6: Degree distribution of all the refined models

4.5.2 Factors driving the formation of links

The results of the formation model are presented in Table 4.3. The model with all links (initial) and the model excluding the 36-month links (refined) are very similar and do not contain conflicting results. The initial model, however, contains more significant coefficients and therefore serves as a basis for the interpretation.

At the node level, *INCR URBAN*, and *URBAN* are significantly negative in Model 2, which indicates that in the second period (2002-2005), rural organizations participate in more joint projects than urban ones. The variable is insignificant in the other models. The results are not in line with Hypothesis 1, which suggests urban organizations being more likely to form links due to urbanization externalities. We suspect an effect similar to what Illenberger et al. (2013) find for individuals. Organizations might compensate for the lower accessibility of partners with the higher acceptance of partners in rural areas. Alternatively, after the BioRegio initiative ended in 2005 (see, e.g., Dohse, 2000), support became less focused on urban regions, and rural regions gained importance in subsidization schemes. In any case, Hypothesis 1 is not confirmed, as organizations in urban regions are not more actively engaging in subsidized R&D collaboration than rural organizations.

SIZE1 and *SIZE2* obtain significantly positive coefficients in Model 1 and Model 2, respectively. Accordingly, medium-sized and large-sized firms have higher probabilities of link establishment in comparison to small firms (fewer than 50 employees). This fits with our line of argumentation in Section 4.2.1 regarding larger firms having more capabilities and opportunities to establish links. Our findings are in line with the results of Tether (2002), who argues that larger firms might benefit from their size in two ways: First, they are more attractive for cooperation partners (e.g., universities), and, second, they might force their suppliers into cooperation projects.

The coefficient of EAST is significantly positive in Model 1. This supports the findings of Cantner and Meder (2008) – specifically, that East German organizations are more active in subsidized R&D-cooperation, which corresponds to the idea of a policy's stronger support for these regions.

At the dyad level, we found that *COG PROX* was significantly positive in all models. Organizations operating in the same subfields of biotechnology are more inclined to conduct joint R&D. Accordingly, Hypothesis 2 is confirmed, and our results add to the findings of Nooteboom et al. (2007) and Balland et al. (2013), showing that cognitive proximity is an important driver of R&D network formation.

In addition to cognitive proximity, geographic proximity also plays a significant role in the formation of R&D cooperation. *GEO PROX* obtained a significant coefficient in the second refined model but remained insignificant in the first and third models[16]. Thus, in the second period, organizations tend to work together with partners located nearby, which supports Hypothesis 2.

[16] Including *GEO PROX* in Models 1 and 3 led to degenerated results. Thus, we decided to exclude it. Nevertheless, degeneracy itself is an interesting topic and needs further research.

Table 4.3: Results of the two-mode STERGM, formation.

	Initial models¥			Refined models¥		
	Model 1 1998 – 2001 2002 – 2005	**Model 2** 2002 – 2005 2006 – 2009	**Model 3** 2006 – 2009 2010 - 2013	**Model 1** 1998 – 2001 2002 – 2005	**Model 2** 2002 – 2005 2006 – 2009	**Model 3** 2006 – 2009 2010 – 2013
Variables	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
Node level						
MEDICINE (base: INDUSTRIAL)	0.514 (0.392)	-0.209 (0.238)	-0.128 (0.365)	-0.0005 (0.663)	-0.298 (0.417)	-0.165 (0.617)
AGRICULTURE (base: INDUSTRIAL)	1.124*** (0.411)	1.208*** (0.240)	-0.065 (0.399)	0.506 (0.677)	1.306*** (0.425)	-0.517 (0.726)
UNI (base: FIRM)	0.003 (0.286)	-0.068 (0.185)	-0.183 (0.302)	0.159 (0.512)	0.468 (0.326)	0.152 (0.549)
RESEARCH INST (base: FIRM)	0.658** (0.273)	-0.029 (0.185)	-0.013 (0.327)	0.646 (0.486)	0.237 (0.351)	-0.034 (0.675)
SIZE1 (base: SIZE0)	0.349 (0.269)	0.543*** (0.183)	0.588** (0.293)	-0.013 (0.471)	0.340 (0.331)	0.056 (0.586)
SIZE2 (base: SIZE0)	0.896*** (0.295)	0.795*** (0.207)	-0.065 (0.378)	0.506 (0.512)	1.084*** (0.377)	-1.073 (1.096)
EAST (base: WEST)	0.565* (0.308)	-0.317 (0.214)	-0.611 (0.421)	1.263** (0.518)	0.226 (0.387)	-0.320 (0.809)
INCR URBAN (base: rural)	0.269 (0.361)	-0.483** (0.230)	0.052 (0.434)	0.209 (0.572)	-0.933** (0.428)	-0.196 (0.774)
URBAN (base: rural)	0.141 (0.359)	-0.777*** (0.222)	-0.269 (0.411)	0.492 (0.639)	-0.954** (0.409)	-0.929 (0.929)
Dyad level						
COG PROX	0.226*** (0.045)	0.185*** (0.032)	0.290*** (0.071)	0.344*** (0.073)	0.422*** (0.069)	0.389*** (0.149)
INST PROX	0.115* (0.062)	0.021 (0.049)	0.063 (0.096)	0.326*** (0.074)	0.216*** (0.081)	0.303 (0.184)
GEO PROX					0.782*** (0.153)	
Structural level						
EDGES	-7.166*** (0.629)	-6.444*** (0.366)	-7.202*** (0.610)	-6.804*** (1.103)	-4.691*** (0.745)	-6.264*** (1.119)
GWDSP, 0.3, fix	-0.262*** (0.052)	-0.130*** (0.031)	-0.294*** (0.065)	-0.579*** (0.099)	-0.837*** (0.101)	-0.551 (0.149)
$GWDEGREE_{B1}$, 0.5, fix	4.187*** (0.451)	3.663*** (0.322)	3.748*** (0.363)	6.278*** (0.856)	4.998*** (0.561)	3.462*** (0.686)
B2DEG3	1.721*** (0.236)	1.674*** (0.166)	2.315*** (0.266)	1.431*** (0.314)	1.274*** (0.236)	1.817*** (0.413)
Null deviance:	57,320 on 41,348 df	179,014 on 129,131 df	168,117 on 121,271df	23,338 on 16,385 df	73,062 on 52,703 df	46,680 on 3,367 df
Residual Deviance:	2,533 on 41,331 df	4,996 on 129,114 df	881 on 121,254 df	1,064 on 16,818 df	2,147 on 52,685 df	-76 on 33,656 df
AIC	2,567	5,030	915	1,064	2,183	-42
BIC	2,713	5,196	1,081	1,230	2,343	100

Institutional proximity (*INST PROX*) is significantly positive in the first formation model, suggesting that organizations with the same institutional background are more likely to work together. Hypothesis 3 is thereby confirmed. Due to less uncertainty regarding partner goals and behavior, organizations tend to select cooperation partners from the same institutional background (Ponds et al., 2007).

Only one of the findings on variables at the structural level is in line with our expectations. All other factors excluded, the variable *EDGES* represents the density of the network and can be interpreted similar to an intercept. As the observed network is the consequence of a social process, it is typically less dense than exponential random networks leading to the negative coefficient of *EDGES* (Lufin Varas, 2007).

Unexpectedly, *GWDSP* was significantly negative in all of the models. This contradicts the multi-connectivity proposition of organizations' tendency to connect through several ways in order to decrease link dependencies. In our case, organizations rarely engaged with the same organizations in multiple subsidized R&D research projects, which appears to be a valid, but still unexpected, strategy to maximize learning and inter-organizational knowledge diffusion. While a potential explanation might be a policy, penalizing collaborations of the same organizations in its subsidization programs, we are not aware of such a rule.

$GWDEGREE_{B1}$'s coefficient gained a significantly sign; however, its sign is positive, which contradicts the preferential attachment process (Hunter, 2007): Organizations are less likely to gain additional links when they are already well connected. We clearly must reject Hypothesis 4 with respect to the link formation model. There are three potential reasons for this: Firstly, organizations are limited in their collaboration capacities, thus implying that they constantly face a trade-off between maintaining and acquiring new links through projects. Similarly, they might not have the capacity or willingness to apply to multiple subsidization programs within the same time period. Secondly, subsidization programs are more focused, and there is only a limited overlap between organizations' activity portfolios and support programs. Thirdly, a policy might favor subsidizing a broad range of organizations and therefore penalizes organizations already active in numerous projects.

4.5.3 The dissolution models

As expected (see Section 4.5.1), we found fewer significant coefficients for the dissolution models (see Table 4.4). We believe that this is due to the relatively low variance in link duration, which is strongly constrained by the design of the underlying policies (4.5.1). Nevertheless, as argued in Sections 4.4.1 and 4.5.1, significant results are still possible and interesting.

The coefficient of *RESEARCH INST* is significant and negative in Model 3. This finding implies that research institutes are either leaving projects earlier (unlikely) or initially opting for shorter projects (more likely) than firms. As research institutes are inclined to exchange knowledge with diverse sources (Ponds et al., 2007), shorter collaboration appear to be more attractive to these organizations. This also allows for the establishment of a diverse network of collaboration partners and for the maximizing of access to knowledge from different subfields. The same argument can be brought forward regarding universities. It might also be the case, however, that both types of organizations relate their R&D projects to the completion of PhD theses (which usually require about three years) and therefore target the 36-month projects. In the case of universities, some support for this can be found in period 3, in which the coefficient is positively significant. In other words, once the 36-month projects are excluded (which are likely to relate to PhD projects), universities are less likely to be engaged in shorter projects and collaboration.

In the Model 3, *INCR URBAN* is significantly negative, meaning that organizations located in urban areas are more likely to dissolve links in comparison to organizations in rural areas. Again, there might be multiple explanations for this. Organizations in urban regions are known to have a large selection of (nearby) potential collaboration partners, which organizations in rural regions lack (Meyer-Krahmer, 1985). Accordingly, they might be more interested in shorter projects in order to exploit and thereby make use of this potential. Organizations in rural regions might also be less attractive collaboration partners because of lower reachability, less prestigious names, etc. This lack of attractiveness has to be compensated by larger subsidies, i.e. larger and longer R&D projects. Additionally, organizations in urban and rural regions might have different technology foci. Shorter projects are more attractive for organizations seeking to remain at the technology frontier, which implies making quick progress and constantly exploring new developments on a short-term basis. However, organizations in rural regions are less likely to be active in the most recent and most complex technologies (Hägerstrand, 1967; Balland and Rigby, 2017). Hence, shorter projects are not as attractive for them, thus leading to lower link dissolution probabilities. Future research should more thoroughly address this issue, such as by applying qualitative methods.

Table 4.4: Results of the two-mode STERGM, dissolution.

	Initial models[¥]			Refined models[¥]		
	Model 1 1998 – 2001 2002 – 2005	**Model 2** 2002 – 2005 2006 – 2009	**Model 3** 2006 – 2009 2010 - 2013	**Model 1** 1998 – 2001 2002 – 2005	**Model 2** 2002 – 2005 2006 – 2009	**Model 3** 2006 – 2009 2010 - 2013
Variables	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
Node level						
MEDICINE	0.054	- 0.103	-0.880**	0.417	-0.393	-0.433
(base: INDUSTRIAL)	(1.671)	(0.407)	(0.310)	(1.151)	(0.515)	(0.403)
AGRICULTURE	0.454	-0.129	-0.167	0.706	-0.090	-0.578
(base: INDUSTRIAL)	(1.665)	(0.419)	(0.328)	(1.338)	(0.523)	(0.419)
UNI	0.203	-0.692**	0.136	0.568	-0.370	0.688**
(base: FIRM)	(1.177)	(0.289)	(0.259)	(0.998)	(0.359)	(0.344)
RESEARCH INST	-0.157	0.175	-0.817***	1.387	0.257	-1.570***
(base: FIRM)	(1.043)	(0.298)	(0.271)	(1.154)	(0.408)	(0.434)
SIZE1	-0.987	0.262	-0.196	-0.758	-0.356	-0.342
(base: SIZE0)	(1.102)	(0.293)	(0.269)	(0.981)	(0.372)	(0.358)
SIZE2	-0.485	-0.036	-0.388	-0.884	-0.042	0.078
(base: SIZE0)	(1.116)	(0.317)	(0.295)	(1.050)	(0.429)	(0.295)
EAST	-1.354	0.197	-0.102	-1.747	0.027	0.201
(base: WEST)	(1.063)	(0.349)	(0.316)	(1.056)	(0.442)	(0.436)
INCR URBAN	-0.664	0.604	-0.842**	-0.859	-0.541	-2.040***
(base: RURAL)	(1.399)	(0.391)	(0.452)	(1.222)	(0.529)	(0.523)
URBAN	-0.788	0.197	-0.476	-0.689	-0.713	-1.154**
(base: RURAL)	(1.455)	(0.402)	(0.336)	(1.281)	(0.512)	(0.485)
Dyad level						
COG PROX	0.03	0.039	-0.129	-0.259	-0.161	-0.195
	(0.349)	(0.102)	(0.083)	(0.361)	(0.170)	(0.161)
INST PROX	0.109	-0.248*	-0.152	0.070	0.086	-0.415**
	(0.366)	(0.126)	(0.095)	(0.348)	(0.149)	(0.196)
GEO PROX	0.32	0.148	0.335	0.128	0.213	0.375
	(0.848)	(0.231)	(0.257)	(0.840)	(0.278)	(0.44)
Structural level						
EDGES	5.295**	1.943***	2.675***	2.606	3.261**	3.537***
	(2.638)	(0.715)	(0.582)	(2.631)	(1.151)	(0.992)
GWDSP, 0.3, fix	-0.003	-0.029	0.115**	0.354	-0.302**	0.187
	(0.267)	(0.074)	(0.051)	(0.375)	(0.139)	(0.164)
$GWDEGREE_{B1}$, 0.3, fix	-0.384	-0.446	-0.465	0.137	-0.405	-0.459
	(1.127)	(0.316)	(0.283)	(1.488)	(0.579)	(0.505)
$GWDEGREE_{B2}$, 0.3, fix	-6.800***	-4.865***	-5.208***	-3.978**	-5.398***	-5.911***
	(1.860)	(0.540)	(0.445)	(1.621)	(0.862)	(0.161)
Null deviance:	327 on 236 df	650 on 469 df	952 on 687 df	228.7 on 165 df	400 on 289 df	952 on 687 df
Residual deviance:	48 on 218 df	467 on 451 df	568 on 669 df	47 on 147df	-72,559 on 271 df	568 on 669 df
AIC:	84	503	604	83	-72,523	604
BIC:	146	578	686	139	-72,457	686

NAs have been excluded.
* significant at the 90% level, ** significant at 95% level, *** significant at 99 % level
[¥] Initial models including the whole network, refined models without links of 36-months length.

We also determined that the dyad-level variable *INST PROX* was negatively significant in the second model. This contradicts Hypothesis 6, in which we argued that institutional proximity is unlikely to influence link dissolution. Here, the negative sign inclines partnerships between profit and non-profit organizations to last longer than between profit and profit organizations and non-profit and non-profit organizations. A straightforward explanation is that projects involving partners with different institutional backgrounds require more time (and hence apply for longer projects) than partners operating within the same institutional framework (Boschma, 2005).

At the structural level, *EDGES* and $GWDEGREE_{B2}$ were highly significant in all the models. *EDGES* is interpreted as in the formation model with its significantly positive coefficient pointed toward higher network density than in a random network. The effect of preferential attachment is also present in the duration of links. The significantly positive coefficient of $GWDEGREE_{B2}$ implies links established between new organizations and projects that are already well embedded in the network are less persistent. We interpret this as being primarily a technical effect. Projects and organization in the network that hold central positions do so because they are participating in large projects. Note that we established earlier that few organizations are active in multiple projects at the same time. Hence, when the project is completed, they will lose most if not all their links at the same time. This number will naturally be larger than in case of less central organizations and projects (because otherwise their centrality would not be lower). Accordingly, prominence in the network caused by participation in larger projects (in terms of the number of participants) tends to imply larger dissolution rates of links. The results for the other structural network variable *GWDSP* are inconclusive as its coefficient alters between the models. Again, we must reject Hypothesis 4, network structural effects do not relate in the expected way to the evolution of the network. This is most likely, partly caused by the endogenous dissolution processes, which are strongly impacted by the (externally fixed) conditions of the support programs.

4.6 Conclusion

In urban studies and related fields, dynamic network analysis has become a crucial tool to understand the evolution of different types of networks in time and space. In particular, studies analyzing spatial knowledge networks have increasingly relied on dynamic network analysis (Glückler, 2007; Boschma and Martin, 2010; Ducruet and Beauguitte, 2014; Glückler and Doreian, 2016). Interestingly, most existing studies have thereby focused on the formation of links. However, as Glückler (2007) put forward, network evolution is a twofold procedure that "should be conceived as the result of endogenous mechanisms of network formation and dissolution" (Ibid: p. 627). Accordingly, in order to fully understand the evolution of spatial networks, both processes need to be considered in empirical investigations.

The chapter contributes to the literature by discussing the separable temporal exponential random graph model (STERGM) as a novel and interesting tool in this context. We demonstrate its use for the analysis of the evolution of spatial (knowledge) networks by presenting a case study on the (subsidized) R&D collaboration network of the German biotechnology industry. In particular, we highlight the STERGM's capacity to directly analyze two-mode networks, which avoids the sometimes questionable one-mode projection (see also Liu et al. 2015). In addition to the dynamic analysis and the possibility of disentangling formation and dissolution, this feature was frequently argued to be the primary benefit of using stochastic actor-oriented models (see, e.g., Liu et al., 2013).

Besides advocating the use of the STERGM, the chapter also aimed at exploring the roles played by location (urban – rural) and different types of proximities (cognitive, institutional, geographic) for the formation and dissolution of spatial knowledge links, with the latter having received little attention in the past.

Table 4.5 provides an overview of the main results. Overall, the results of the formation models are in line with the theoretical expectations. Interestingly, the same cannot be said for the dissolution models. In these cases, we were not able to find solid evidence for location and the proximities to strongly impact link dissolution.

Table 4.5: Summary of the main results

		Formation		Dissolution	
Hypothesis	**Variable**	**Result**	**Supporting hypothesis?**	**Result**	**Supporting hypothesis?**
Node level					
H1	URBAN	Negative relationship	No	Insignificant	No
Dyad level					
H2	COG PROX	Positive relationship	Yes	Insignificant	No
H2	GEO PROX	Positive relationship	Yes	Insignificant	No
H3	INST PROX	Positive relationship	Yes	Negative relationship	No

However, we observed that these factors seem to vary in their influence on formation and dissolution. If we expand our view beyond our relatively narrow hypotheses, we find confirmation for variations in the relative importance of factors for link formation and dissolution; some factors are more crucial for the formation while others impact link dissolution to a greater extent. For instance, institutional proximity, i.e. whether organizations cooperate within the same (university, applied research, or profit) framework, makes link formation more likely. At the same time, it also facilitates link dissolution. Accordingly, simply inferring from knowledge on formation processes on dissolution dynamics is invalid, and we need to analyze both processes separately.

Our results also show that factors' influence on network evolution is not time-invariant but is instead conditional on the current framework in which an industry operates. While previous studies argued for the relevance of industry life-cycle phase and thereby endogenous conditions (Balland et al., 2013; Ter Wal, 2014), our analysis (due to the nature of the employed data) highlights the relevance of external circumstances — in this case, variations in the R&D policy.

As is typical for empirical studies, this case study used for demonstrating the applicability of STERGM is subject to certain limitations. First, the STERGM was only recently developed, which implies some shortcomings that will certainly be addressed in the future. Currently, continuous variables at the dyad level are difficult to implement. This is particularly relevant in the context of spatial networks as geographic proximity is usually modeled in a continuous way. As of now, researchers working with the STERGM need to work with categorical definitions. Second, the robustness of the simulated networks, i.e. of the model converges, depends on a variety of factors that are hard to isolate (e.g., network size and continuous variables such as the amount of funding). This implies considerable difficulties in terms of finding the best-fitting model. Third, the discrepancy between methodological possibilities and data availability is the most apparent shortcoming of our study. This chapter highlights and promotes the STERGM's feature of disentangling link formation and dissolution processes. However, when looking at the most commonly used data for constructing spatial networks (such as that in the present chapter), it turns out that most of the data encounter the same issues: either there is no (precise) available information on the duration of links (e.g., patent data, co-authorship data) or, if this information does exist, it might be subject to external conditions (e.g., relations established on the basis of the subsidization of joint R&D projects). Accordingly, while the methodological precision and possibilities to explore (spatial) network evolution continuously increase, the same cannot necessarily be said about the available data. Hence, researchers need to be aware of the gap existing regarding the methodological possibilities and what can actual been done with the data at hand. The opportunity to explore longitudinal two-mode network data with dynamic network analyses is hence a step in the right direction as it moves the methodological side closer to the type of data available. Nevertheless, we clearly pledge for more efforts to be directed toward the collection of data on link dissolution, as otherwise our understanding of knowledge network evolution will remain constrained.

Despite these shortcomings, some policy implications can be derived from the present study. Firstly, our results indicate that institutional proximity is still an important determinant of link formation. Given the wide belief in the necessity to involve heterogeneous sets of actors in R&D projects and that spillovers between the non-profit and profit sectors are to be increased (see, e.g., triple helix literature (Etzkowitz and Leydesdorff, 2000)), these goals are not yet visible in our results. Profit organizations still seem to prefer to work with other profit organizations, and non-profit organizations are more frequently engaged with other non-profit organizations.

Secondly, as with other related studies, we found that proximities are important drivers of subsidized network formation. One can argue that these represent the “natural” way in which networks evolve without external influences. This is confirmed in many analyses on non-subsidized knowledge networks (Glückler, 2010; Balland et al., 2013; Ter Wal, 2014). Hence, networks influenced by a policy and those that are not influenced by such evolve in the same manner — i.e., they have the same factors driving their evolution. If this is the case, it may lead one to wonder why policy is providing subsidies for collaboration in the first place. When a policy supports the same kind of interactions that evolve independently of it, in the best of all cases, it merely increases the general magnitude of collaborations. However, it does not impact their structural composition. This particularly concerns cognitive proximity, which makes the establishment of projects generating significant novelty less likely (Boschma, 2005; Nooteboom et al., 2007). In this respect, this chapter calls for a reconfiguration of the R&D subsidization policy.

Appendix

4.A1 Network characteristics

Table A4.1: Network characteristics

Model	Networks	All nodes	Project nodes	Organizational nodes	Links	Density
Initial networks						
1	No. 1	481	113	368	236	0.0020
	No. 2a	481	113	368	494	0.0040
2	No. 2b	816	216	600	494	0.0014
	No. 3a	816	216	600	687	0.0021
3	No. 3b	789	211	578	687	0.0022
	No 4	789	211	578	530	0.0017
Refined networks						
4	No. 1	318	68	250	165	0.0032
	No. 2a	318	68	250	295	0.0059
5	No. 2b	542	128	414	289	0.0019
	No. 3a	542	128	414	353	0.0024
6	No. 3b	425	107	318	353	0.0039
	No 4	425	107	318	226	0.0025

4.A2 Variable descriptives

Table A4.2: Variable descriptives (initial network).

Variable	Category	Category Name	Description	Observations
COG PROX	1	MEDICINE	Medicine	272
	2	AGRICULTURE	Agriculture	175
	3	INDUSTRIAL	Industrial processes	86
	4	BIOINFORMATICS	Bioinformatics	25
	NA			94
INST PROX	0	FIRM	Private	191
	1	UNI	Universities	283
	2	REASEARCH INST	Research Institutes	178
SIZE	0	0	< 50 employees	289
	1	1	< 250 employees	134
	2	2	> 250 employees	229
EAST	0	WEST	West German region	548
	1	EAST	East German region	104
GEO PROX	83 different categories		Organization sharing a NUTS 3 region	Max: 30 Min: 1 Average: 3.6
URBAN	0	RURAL	Rural region	98
	1	INCR URBAN	Increasing Urbanization	175
	2	URBAN	Urban region	379

Pulled or pushed? The spatial diffusion of wind energy between local demand and supply

Abstract

This chapter contributes to and connects the literature on spatial innovation diffusion, entrepreneurship, and industry life-cycles by disentangling the relevance of local demand and sup-ply in the adoption of wind energy production. More precisely, we evaluate the strength of local supply-push effects with those of local demand-pull over the course of the evolution of an industry and its main product evolution.

By using Bayesian survival models with time-dependent data of wind turbine deployment and firm foundation for 402 German regions between the years 1970 and 2015, we show that the spatial evolution of the German wind energy industry was more strongly influenced by local demand-pull than local supply-push processes. New producers are found to emerge in proximity to existing local demand for wind turbines. No evidence was found for producers being able to create local demand for their products by pushing the adoption of the technology in their regions.

This chapter is co-authored with Tom Broekel. The PhD candidate is the first author of the article. The paper has been accepted for publication in Industrial and Cooperate Change, 2020.

5.1 Introduction

A growing literature is investigating the emergence and evolution of industries across space and time. The economic geography and regional science literature usually focuses on the role of the (local) supply side in the emergence and evolution of industries, emphasizing agglomeration externalities, path dependence, windows of local opportunity, and related variety (Bergek and Jacobsson, 2003; Boschma and Wenting, 2007; Fornahl et al., 2012). Notably, while the demand side has received much less attention in the field, its relevance has been highlighted in the diffusion of innovation literature originating from Rogers (2003) and (Hägerstrand, 1952, 1965, 1966). More recently, the (primarily sociological) literature on technology transition has underlined demand processes as crucial for the emergence and expansion of industries and their products[17] (Geels, 2004). In particular, in this literature, local demand (in combination with local institutions) is argued to create (market) niches within which new industries and products can grow before facing the "full" competitive forces of non-local markets (Schot and Geels, 2008).

Hence, both lines of argument emphasize the role of local demand and supply processes for the emergence and growth of industries. However, while a substantial empirical literature exists assessing the role of regional supply-side factors for industries' emergence, empirical findings on the importance of local demand are less extensive. In addition to both literature streams paying more attention to the early stage on industries' developments, empirical studies in these fields rarely analyze demand and supply factors side-by-side (Justman, 1994). Accordingly, it is still not well understood how local demand relates to the emergence and concentration of industries, and to what extent industries may themselves contribute to the activation and formation of local demand.

The present chapter contributes to this debate by investigating how industries' evolution is shaped by local supply-side conditions (i.e. manufacturers) and the spatial distribution of demand. In the context of grand societal challenges, such as climate change and the associated transition toward renewable energies, demand becomes particularly important as it enables producers to learn about these new and changed consumer needs as well as shifts in their preferences (Martin et al., 2019). Therefore, this chapter evaluates the (statistical) impact of local wind turbine producers on regional wind turbine installation and the extent to which existing regional wind turbines stimulate the emergence and extension of local wind turbine production. Extending the work at the level of product innovation (e.g. Brem and Voigt, 2009),

[17] For the sake of readability, we stick to the term "product". However, our argumentation is also valid for new technologies.

we compare the contribution of local supply-push and local demand-pull processes to the emergence and evolution of the industry over multiple stages of its life-cycle.

For our empirical analysis, we use data on wind turbine deployment and firm foundation for 402 German regions for the years 1970–2015 and employ a Bayesian event-history analysis (Zhou and Hanson, 2018). To analyze both the supply and demand side, we first explain the spatial diffusion of wind turbines considering the location of manufacturers. Second, we investigate the location decisions of manufacturers using the (regionally) existent and future installments of wind turbines as approximations of local demand. Our results show that manufacturers emerge more frequently in places with already existing wind turbine installations. Hence, local supply-push factors, in the form of local niche building, are identified to be less relevant while local demand-pull plays a greater role.

The chapter is structured as follows: Section 5.2 gives an overview of the underlying theoretical arguments, building on the literature of economic geography, innovation, and transition studies. Section 5.3 describes the evolution of the German wind industry. The empirical design of our study is presented in Section 5.4. Section 5.5 shows the results and Section 5.6 concludes with a discussion of the results, shortcomings, and future research.

5.2 The emergence and evolution of industries in time and space

5.2.1 Emergence

To explain the spatial origins of new industries, Scott and Storper (1986) and Storper and Walker (1989) developed the Window of Locational Opportunity (WLO) concept. Here, so-called “trigger events” mark the starting point for the emergence of new industries in specific locations. The industry’s initial locations are distributed relatively arbitrarily and unpredictably, as their needs in terms of resources and skills are diverse and distinct from the older existing industries (Boschma and Lambooy, 1999). Consequently, emerging industries are characterized by relatively high degrees of freedom in terms of location. In later extensions of the concept, the assumption of the randomness of locations was revised with greater importance assigned to regional conditions (Boschma and Lambooy, 1999; Fornahl et al., 2012). In particular, scholars have argued that the likelihood of new industries emerging in regions grows when related industries are already present (Boschma and Frenken, 2011). Technological relatedness thereby refers to a certain but not complete cognitive overlap, similarity, and complementarity of core technologies, and potential shared development history (Frenken et al., 2007). The presence of related industries ensures the availability of required skills, human capital, infrastructure, and collaboration partners. The process of industry emergence and expansion on the basis of regional technologically-related industries is referred to as regional branching (Boschma and Frenken, 2011) and manifests itself through various diversification mechanisms (Asheim et al., 2011; Buenstorf et al., 2015). For example, the diversification of existing firms from related industries has been highlighted as a catalyst for the first regional entry of new industries (Helfat and Lieberman, 2002, Klitkou and Coenen 2013). Regional branching processes are also fueled by spin-off activities with entrepreneurs from related industries (Boschma and Wenting, 2007; Klepper, 2007). Both mechanisms, related diversification and spin-offs, have a strong regional dimension in that firms tend to establish new activities near to existing operations, and spin-offs tend to be located close to their parent company. Other processes fueling regional branching are spatially limited labor mobility and knowledge diffusion in social networks, which are also more intense within relatively small-scale areas and among related industries (Asheim et al., 2011).

5.2.2 Concentration

While some regions manage to become the initial locations of new industries, they do not necessarily benefit from their subsequent growth as industrial concentration processes frequently lead to industrial agglomeration in only a few regions. But which regional characteristics favor such concentration processes? The literature addresses this question among others with concepts such as industrial districts (Marshall, 1920), Italian industrial districts (Pyke et al., 1990), clusters (Porter, 1998), and innovative milieus (Camagni, 1995).[18] In evolutionary economic geography, spin-offs and agglomeration externalities are emphasized as explanatory mechanisms (Arthur, 1994; Klepper, 2006). Spin-off processes are transmission channels in which routines and knowledge diffuse from parent organizations to new enterprises. The mechanism has features of a self-reinforcing, snowball-like process, as the likelihood of further spin-offs depends on the number of existing firms in a region. The importance of spin-off dynamics in the emergence of regional industry clusters has been demonstrated by numerous examples, such as the information and communication technology industry in Silicon Valley (Saxenian, 1994) and the automotive industry in Detroit (Klepper, 2007).

In addition to spin-off processes, agglomeration externalities may play significant roles in regional industry concentrations. These externalities emerge from the co-localization of economic actors (Neffke et al., 2011) and represent the spatial connotation of increasing returns (Krugman, 1991). It is customary to differentiate between Jacobs and Marshall externalities,

[18] It is beyond the scope of the present chapter to present and discuss these concepts. An overview can be found in (Brenner and Mühlig, 2013: 482 ff.).

where the first refers to externalities resulting from the spatial concentration of economic actors in different activities and the latter relates to effects resulting from the agglomeration of firms in the same sector (Boschma and Wenting, 2007). In addition, urban regions may offer advantages to young industries, as their size and greater economic diversity make firms more able to find generic resources, among which are human capital, services, and infrastructure (Hoover and Vernon, 1962). The discussion of these types of externalities has recently been extended to include a more dynamic approach, which considers industries not only being exposed to regional conditions but also being able to contribute to their development. This is taken up in the concept of related variety, which highlights concentration-promoting externalities that often emerge from the agglomeration of existing, related industries (Boschma and Wenting, 2007; Boschma and Frenken, 2011). It highlights mutual adaptation processes between new industries and their regional environment. For example, research and development (R&D) investments generate industry-specific knowledge and employees gain industry-specific skills through on-the-job learning. As an industry becomes more established, the specific resources that are created increase in importance and may even reach a critical mass such that increasing demand for them leads to the emergence of an efficient "local production environment" (Boschma and Lambooy, 1999). This may stimulate regional branching processes or attract further firms from related industries and, hence, may give rise to self-reinforcing processes according to the principle of cumulative causation (Myrdal, 1957). From studies covering the multiple development phases of industries, it has been observed that the relevance of the different processes and factors (i.e. branching, related variety, spin-offs, and agglomeration externalities) changes. Moreover, the effect of spin-off processes with an industry-specific background is less relevant for industry concentration in the development phase of an industry due to the (still) low potential of the parent companies—at this stage, related industries are more significant (Boschma and Wenting, 2007). A large number of empirical studies confirm these processes, showing the path-dependent emergence and development of industries in space (e.g. Balland et al., 2013; Breul et al., 2015). For instance, Neffke et al. (2011) show that young industries benefit from Jacobs externalities, whereas more mature industries profit more strongly from Marshal externalities.

5.2.3 The creation of local technological niches

The literature reviewed in the previous subsection noticeably pays more attention to “supply” or “push” dynamics. That is, the likelihood of industry emergence and spatial concentration are primarily explained based on the availability of “input factors” such as human capital, infrastructure, resources, knowledge, and the presence of related competencies. While not absent, local demand and supply factors have received less explicit attention in theoretical arguments as well as in empirical studies. For instance, the presence of local customers, for example, in the form of related industries positioned at later stages of the value chain, is clearly acknowledged in the discussions on related diversification and urbanization externalities. Yet, few discussions are found in this literature stream on the presence of potential end-users in regions and how these may contribute to the emergence of an industry or its spatial concentration.

In the neoclassical models of Weber (1909) and later Myrdal (1957), it is argued that demand is an important factor for the location decision of firms and hence for the spatial distribution of industries. However, both argue that demand is an indirect factor for location decisions as demand has a positive effect on supply-side factors. More recently, a more dynamic view of the demand side has been put forward in transition studies (Geels, 2004; Schot and Geels, 2008). Transition studies build upon the notion of institutional embeddedness in socio-technical systems (STS). These include three elements: production, diffusion, and the use of technology including both supply and demand (Geels, 2004). These subsystems shape and are shaped by the actions of actors on three different but interrelated levels: the “technological niche,” “the socio-technical regime,” and “landscape” (Geels and Schot, 2007). Geography enters by combining the socio-technical dimension with that of socio-spatial embeddedness (Truffer and Coenen, 2012). More precisely, it takes a “spatially informed, co-evolutionary transition model” (Ibid: 11) that considers the evolution of new industrial niches as an asymmetric process of regional development. Hence, when analyzing the emergence of new technological niches, not only the regime and technological landscape need to be looked at, but the regional context in which the niches are built up as well.

When new products are pushed into the market (supply-push or technology push) (Brem and Voigt, 2009), the new technology is usually superior to existing products in some way and therefore has the potential to create its own demand. However, few customers are willing to buy and test the product. Only “innovators” or “early adopters” are willing to take a risk by spending money on unknown products (Rogers, 2003: 22). Consequently, demand is usually too low (e.g., due to high production or utilization costs) for new products to sustain themselves in fully competitive environments (Geels and Schot, 2007). The emergence of “technological niches” is therefore critical for their survival. These “technological niches” are a kind of protected area and include a network of supportive actors (Schot and Geels, 2008). Jacobsson and Johnson (2000) called these actors partly “prime movers”, who contribute to niches by raising awareness, undertaking investments, providing legitimacy, and facilitating the diffusion of new products. In particular, inventors and producers of new products play a decisive role in the formation of those supportive networks and niches as they “shape the selection process itself by setting up special programs in R&D settings or demonstration projects” (Schot and Geels, 2008: 539).

Support networks tend to be place specific, as geographical proximity between contributing actors facilitates interaction and coordination possibilities, which in turn help the network to form and grow. If successful, a support network will expand in size (in terms of actors and space) and establish its own local institutions, routines, and dynamics (Coenen et al., 2012) after which the niches might reach the stage of a regime (Essletzbichler, 2012). Besides supportive networks, geographic proximity is also important for the local demand side because prime movers frequently create new niche markets that are geographically separated from the main market by establishing “local practices” (Geels and Deuten, 2006). Hence, markets for new products are likely to be located in geographic vicinity to their producers, as these may have created the markets themselves. Moreover, lacking established marketing and distribution channels, the diffusion of knowledge about these new products tends to be hampered by geographic distance (Hägerstrand, 1965). In some instances, this may be strengthened by new products’ limited transportability when product-specific transportation infrastructure is required and needs to be established first. Hence, in the early stages, the possibilities of serving geographically large-scale markets may be significantly reduced. Over time, infrastructure and distribution, as well as marketing channels, will be built up, and consumers at larger geographic distances can be served. As a prerequisite, the establishment of local technological niches including local demand is essential. In this case, it is the emergence of new products that create their own local demand and thereby shape the spatial distribution of industries. Put more bluntly, these arguments suggest that the spatial distribution of supply shapes the subsequently developed spatial distribution of demand. Noticeably, the creation or activation of local demand by the supply side is particularly relevant in the early, emergence stage of a product or industry. However, another relationship between demand and supply may be of greater importance in other stages.

5.2.4 Regional demand as a pull-factor

According to the STS literature, demand rises when a certain socio-technical regime is not in a state of equilibrium, in other words, the current technology does not fit or satisfy user preferences. Such an imbalance might originate from the socio-technical landscape (e.g., new dominant lifestyles) which can “modify the direction of development paths and innovation activities” (Geels and Schot, 2007: 406). According to Essletzbichler (2012), these landscapes are “ […] multiple selection environments operating on various spatial scales” (p. 798). Part of the variations in landscapes are regional differences in demand and consumer preferences. Understanding these preferences is important for the success of a product and its producer. The producer may learn about consumer preferences and their changes through interactive and collaborative learning with consumers and users (Martin et al., 2019). This is particularly the case for industries in which innovation processes are strongly linked to doing, using, and interacting (DUI) (Jensen et al., 2007). This applies to the early phases of the development of the wind industry when synthetic knowledge was combined with experience-based skills and crafts. This "bricolage" process is described in detail by Garud and Karnoe (2003).

For instance, the size of regions may matter in this context. If potential early adopters make up a small share of the population, the absolute initial regional demand for products will scale with the size of the population. Moreover, given the smaller distances between potential consumers, information about a new product will diffuse faster within the region. Put differently, larger local technological niches (at least from a demand perspective) are more likely to be formed in urban areas than in rural ones, which is in line with the work of Hägerstrand (1965). However, size is not the only regional characteristic that may matter in this context**.** Regions also differ in accumulated experiences and the presence of tacit knowledge with respect to products, as well as in actors’ propensity share this knowledge among producers (Martin et al., 2019). In addition, existing infrastructure, natural resource endowments, social and cultural capital, wealth and culture may also translate into regional differences in demand, making specific regions more attractive for producers.

Being proximate to demand is particularly attractive when transaction and transportation costs are significant. In such cases, firms may choose to locate close to the demand to maximize profit and, hence, demand may “pull” firms to specific locations (Weber, 1909). However, transportation costs have been decreasing on average in recent decades for most products, suggesting that this argument might have lost significance. However, this is not true for all goods. If trucks and heavy-duty transportation are necessary, for example, if the buyer is located in an inaccessible location or if the transport infrastructure is inappropriate, transportation costs may still be of relevance (Ashwill, 2003). Additionally, if components are fragile and upheaval might damage the product, transportation costs tend to rise as well (Kammer, 2011). Accordingly, for some goods and locations, transportation costs are still highly relevant. In these cases, it is economically more favorable for manufacturers to produce in geographical vicinity to demand.

Other factors that can influence the location decision process of firms are the public sector and policy. First, these may create additional demand, such as through public tenders or procurement (Edler and Georghiou, 2007). Second, by offering tax reductions and subsidies, policymakers can support the development of specific industries. For instance, public tenders that are combined with “local-content requirements” are direct monetary effects that shape the location decisions of manufacturers. These have been successfully used by the provincial government of Quebec, Canada when setting up a 1,000 MW wind farm (Lewis and Wiser, 2007), with the winning company General Electric (USA) establishing three manufacturing facilities in Canada as a result. Naturally, public policy is especially likely to support industries in this way when they are associated with providing solutions for societal issues such as climate change and sustainability.

Another process that may lead firms to “follow” demand can be found in the entrepreneurship literature. Some entrepreneurs set up their company to satisfy their own demand. That is, driven by an unsatisfied personal need, inventors initiate the production or development of new products. If the innovation is adopted by further actors, such as family or friends, the inventor might become aware of its business potential and eventually found a company. Shah and Tripsas (2007) call this “user entrepreneurship”. Entrepreneurs frequently set up their companies close to their home (Boschma and Martin, 2010). Accordingly, it can be argued that it is the initial (individual) demand which decides the location of industrial emergence.

Previously, we have presented arguments for the supply side being able to impact the spatial distribution of demand and thereby being relatively more impactful on the spatial distribution of an industry in its emergence phase. When looking in the opposite direction, a less clear statement can be made. Some aspects (such as entrepreneurs satisfying their own demand) are of great relevance in the emergence phase of an industry as well. Transportation costs and regional variations in demand seem to be less specific to a particular stage of an industry’s development. While arguments can be made that the “discovery” of the optimal location in terms of minimizing the distance to demand takes some time and requires the product to have somewhat matured, the issue of inappropriate transport infrastructure is likely to be reduced over time. Hence, it is up to empirical analyses to shed light on these processes.

5.3 The evolution of the German wind industry

To empirically disentangle the contribution of local demand and local supply on the spatial evolution of industries, we take the German wind industry as an example. This is motivated by a number of reasons. First, renewable energy industries are of high importance for policymakers in order to achieve the established climate goals and the new industries promise numerous new jobs (Burton et al., 2011). In addition, the decentralized structure of renewable energy systems generally opens up an interesting research area for geographers (Dewald and Truffer, 2012). Second, the wind industry in its modern form is only a few decades old, which increases the availability of data for the early stages of this industry. Moreover, the geographically-fixed installation of wind turbines and the obligation to report every new plant in Germany until 2015 allow for approximating the geographic (and temporal) distribution of their demand. This is a very appealing feature of this industry because, generally, little information on local demand is available for emerging industries. Third, the production of wind turbines does not require specific natural resources or regional characteristics implying that (in principle) firms in this industry are relatively unconstrained when choosing their location (Kammer, 2011). Moreover, some of the industries that the wind industry is strongly related to and that might spur related diversification processes belong to the mechanical engineering sector (Ibid.). Despite significant regional variations, it can be assumed that basic competences in mechanical engineering are existent in the vast majority of regions in Germany and, hence, the set of potential locations for the industry's emergence is substantial. Accordingly, the WLO covers a significant number of regions. Fourth, the wind industry is characterized by very high transportation costs, which on average amount to 7–10% of total costs (Ashwill, 2003). Should complex or problematic situations lead to higher idle times for trucks, costs can even increase to 20% of the total investment (Kammer, 2011). Proximity to demand is therefore a non-negligible locational advantage.

5.3.1 The rise of the wind energy system

The wind industry started to emerge in the late 1970s in Germany when the first societal and political rethinking of the energy system took place, stimulated by the energy crisis (Simmie et al., 2014). Before then, Germany's energy regime had relied on coal and nuclear power but the crisis led to the first proposals demanding a change to renewables (Jacobsson and Lauber, 2006). R&D expenditures somewhat rose in favor of renewable energy and led to the first development projects, initiating the growth of specialized knowledge (Johnson and Jacobsson, 2003). Within this experimental phase, producers were individuals living in rural areas and building all the necessary infrastructure on their own. Simultaneously, global companies like Boeing or MAN began entering the market (Kammer, 2011).

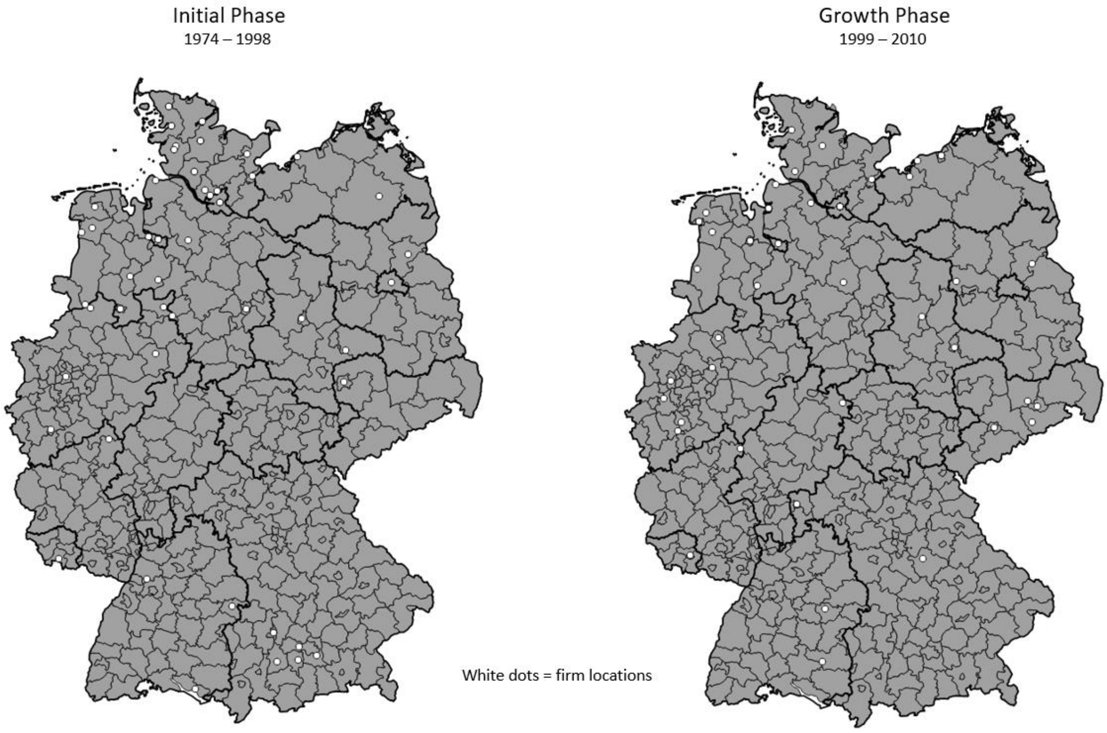

Figure 5.1: Spatial development of the wind energy industry in Germany, 1974 -2010 (Source: own visualization and www.gadm.org, version 2.8, November 2015)

In these first years, new firms scattered over the whole of West Germany (Kammer, 2011). In Lower Saxony (North-Western Germany), a small core of regional producers and suppliers evolved (see Figure 5.1). Former shipbuilders started to manufacture towers and rotor blades for wind energy plants. Gearboxes were produced in South Germany and in the Ruhr Area, as expertise in mechanical engineering was necessary (Kammer, 2011).

From the mid-1980s onwards, events at the socio-technical landscape started to put additional pressure on the dominant regime of energy production: the accident of Chernobyl in 1986, the general climate change debate, and the reunification of Germany (Jacobsson and Lauber, 2006; Kammer, 2011). Policymakers eventually acknowledged changing societal attitudes toward sustainability and renewables, which, amongst others, led to the introduction of the electricity feed-in tariffs law in 1990. This law provided financial certainty for investors by guarantying relatively high feed-in tariffs for a long time period (Kaldellis and Zafirakis, 2011). Moreover, on the product level, the design of three-bladed turbines became the dominant design (Johnson and Jacobsson, 2003). This led to the standardization of production and a cost reduction per kilowatt of 29% between 1990 and 2004 (Kammer, 2011).

During these events, the wind turbine industry experienced its first *take-off* and started to grow and diffuse across the country. For instance, the world’s first fair on wind energy, “Husum Wind”, took place in Lower Saxony in 1989. From 1990 onward, German reunification opened up new manufacturing locations in Eastern Germany (Kammer, 2011 and Figure 5.1) and some firms started to establish their first sales offices outside of the country.

In a typical fashion for an emerging industry, this was a turbulent phase with product standards not yet being defined, consumers still being skeptical, and the well-established energy regime of coal and nuclear power energy producers putting pressure on the new market entrants (Jacobsson and Lauber, 2006; Kammer, 2011). Consequently, many newcomers producing wind turbines quickly exited the market by either filing for bankruptcy or by merging with other firms (Kammer, 2011). For example, in 1989, Vestas Wind Systems bought the Danish producer Danish Wind Technology (Ibid: 153). Crucially, in this phase, such mergers and exits were not part of an industry-wide consolidation process but rather part of the explorative character of innovation and entrepreneurial processes in this phase of the industry’s development.

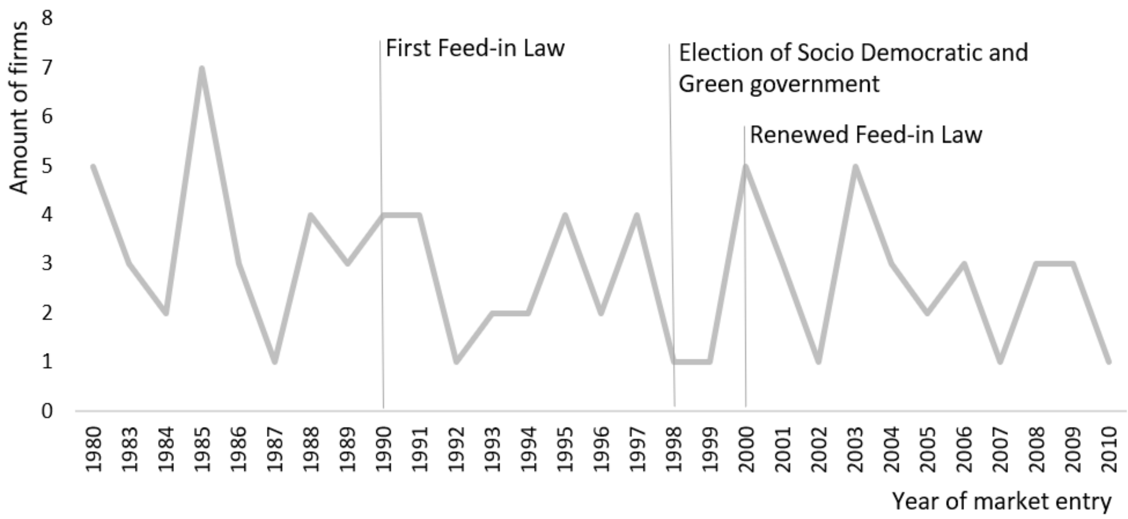

Figure 5.2: Number of firms entering the wind market per year and selected political events

Political support for wind energy and other renewables was significantly strengthened when the Social Democratic/Green coalition gained power in 1998. The coalition initiated market formation programs for renewable energies such as eco-taxes on energy. Further indirect support came from the decision to phase out nuclear power (Jacobsson and Lauber, 2006). Additionally, and most importantly, the nationwide feed-in tariff law was reformed and renewed by adopting the Renewable Energy Sources Act in 2000. This implemented long-term financial support for new renewable energy production and increased investment security (tariff schemes were guaranteed for 20 years), which greatly stimulated the installation of wind turbines (Johnson and Jacobsson, 2003). The industry's growth in this phase led to established manufacturers setting up regional production plants and additional start-ups entering the market (Kammer, 2011 and Figure 5.2).

5.3.2 The wind industry life-cycle

The German wind energy industry has passed through multiple life-cycle phases so far. Based on the work of Klepper (1997), we divide its evolution into three stages: initial, growth, and maturity. This is usually done based on the firm's entry and exit rates. However, as the wind industry is highly subsidized (see Johnson and Jacobsson, 2003), political influence strongly biases these rates (see Figure 5.2). We, therefore, define the industry's life-cycle phase based on the development of its primary product—wind turbines. To abstract from technological specifics and incremental innovation, we capture it by the installed capacity of energy generation from wind.

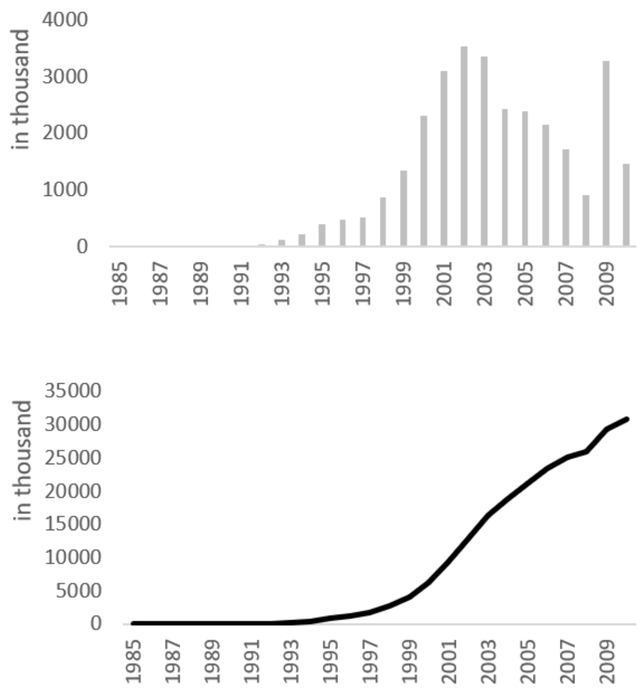

Figure 5.3: Annually installed (top) and cumulated nominal capacity (bottom) in Germany (1995 – 2010) (Source: energy-map.org)

The cumulated installed capacity shows the expected S-Curve of diffusion (see Figure 5.3). Until the end of the 1980s, only several hundred kilowatts were installed every year. From 1989 onward, the typical take-off becomes visible. In 1991, for the first time, more than 10,000 kilowatts were installed. In line with this, Neukirch (2010) assigns the years from 1973/75 to 1991 as the pioneering phase of the industry. Ohlhorst (2009) identifies the years from 1990 to 1995 as the breakthrough years, which was followed by a short decline in growth during the years 1996–1998 (see also Figure 5.3). In line with Klepper (1997), we combine the pioneering phase (the mid-70s to 1991) and the take-off stage (1992–1998) into the *initial* stage.

From 1999 onwards, one million kilowatts were installed yearly (except for during the financial crisis of 2008). Given this qualitative jump in installed capacity, we define the years 1999 to 2010 as the *growth phase* of the industry. Since the yearly installed capacity continued to grow linearly until the end of our observation phase, we cannot identify the industry moving into its maturity phase. This is in line with other studies that argue the industry has not yet reached this stage (Kammer, 2011). Moreover, if we look at the mean values of market entries for both phases, we observe a value of 3.05 for the initial phase and a value of 2.6 for the growth phase. However, the first value is biased by the year 1985 when seven firms entered the market. Excluding this value, we obtain a mean of 2.8, only slightly higher than in the growth phase. For the maturity phase, we would expect a smaller number of market entries (Klepper, 1993).

5.4 Empirical approach

5.4.1 The data at hand

The first data set we use is the wind turbine database (“EEG Anlagenregister”). In Germany, every new renewable energy facility had to be registered in this database by the grid operator until the end of 2015. Most importantly, it includes information about the time of activation, place of deployment, and capacity. Energy-map.org reviewed this information and added the geolocation of all facilities. Amongst others, the database lists all onshore wind turbines starting from 1983 to 2015. On this basis, we observe the time of the first and all subsequent wind turbine installations for every NUTS3 region in Germany. We use this information as an approximation of the regional demand for wind turbines. In particular, we identify which regions where the first to install wind turbines at all.

To model the supply side, we conducted detailed web research in order to identify all wind turbine manufacturers in Germany that existed at some point between 1970 and 2015. We started with the list of manufacturers gathered by Kammer (2011) and extended it by searching on the manufacturers’ own websites and on business registers like *unternehmen24.info*. Additionally, websites such as *Wind-turbine.com* and *www.wind-turbine-models.com* were helpful for acquiring an overview of the industry. In total, we identify 103 manufacturers and collected their date and place of foundation. This includes the location of the firms’ first production facility, as well as the additional production facilities they opened over the years in different regions.

Before presenting the empirical variables constructed on this information in Sections 5.4.3 and 5.4.4, we first introduce our empirical model.

5.4.2 Bayesian spatial survival analysis

The adoption of innovations and the foundation of firms are events taking place at certain moments in time. In the case of innovations, their timing dependents on the innovation itself, the adopter's characteristics, and environmental conditions (Rogers, 2003). With respect to founding a firm, regional characteristics tend to impact the time of establishment (Boschma and Wenting 2007; Sternberg 2003). Therefore, we model the founding of a firm in a location as an event in time, which is related to the prior emergence of its industry at another time and location. To identify the determinants influencing the occurrence of this event, we make use of survival analysis methods.

Survival models originate from medical research and seek to explain how the risk, or hazard, of an event occurring (e.g., death) is conditioned by covariates of theoretical interest (e.g., medication) (Fox and Weisberg, 2011). Recently, these models have been used to study the diffusion of events in time and space. For instance, Darmofal (2009) applies spatial Bayesian survival models to explain the diffusion of political ideas in the United States and Perkins and Neumayer (2005) make use of a Cox proportional hazard model to study whether emerging countries adopt new technologies faster due to smaller investments in prior technologies.

Generally, a survival model consists of the following elements:

$$p_{(t)} = p_0(t) \exp\left(\beta^T x(t)\right)$$

where $p_{(t)}$ is the probability of an event at time *t* (e.g., death), $p_0(t)$ is the exogenous baseline hazard, i.e., the probability of an event occurring independently of any covariates. $x(t)$ is a vector of covariates (e.g., drugs) affecting the baseline hazard and β^T is the corresponding vector of parameters (Perkins and Neumayer, 2005).

A specific reason to use survival analysis instead of standard regression models is the *censoring* inherent to longitudinal data. Censoring defines the possibility that events may lie outside the observation time; called *left censoring* if the event occurs beforehand and *right censoring* when the event takes place after the observational period (see Figure 5.4). In contrast to survival models, standard regression models do not consider censoring and thereby miscalculate the average time it takes for an event to happen (Mills, 2011).

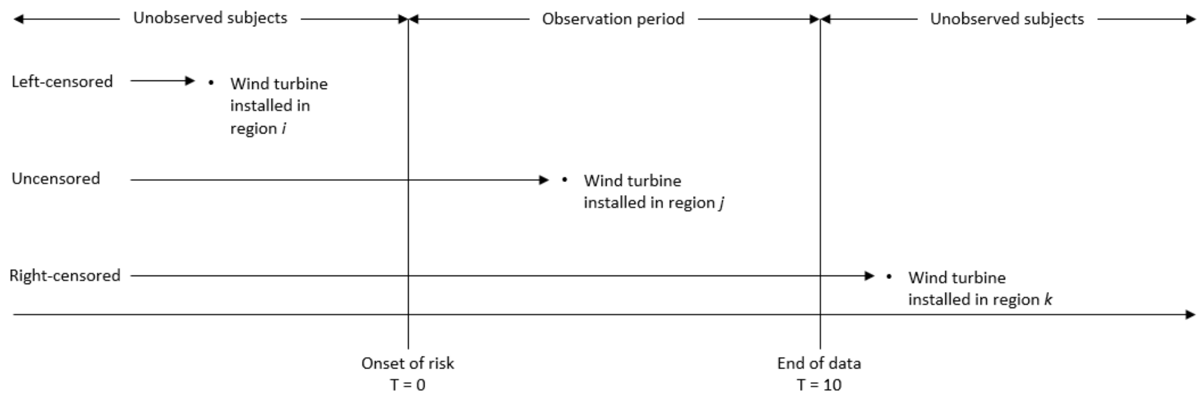

Figure 5.4: Left-censored, uncensored and right-censored subjects

Survival models differ in the distribution of their baseline hazard. For instance, the semi-parametric Cox model does not assume any baseline hazard, i.e., it has no intercept (Cox, 1972). In contrast, the calculation in Weibull or Gompertz models is based on a specific baseline hazard. Thus, researchers must decide between the flexibility of a Cox model or a more precise parametric model, if the assumed baseline is correct (Box-Steffensmeier and Jones, 2004).

Another issue in the use of survival analysis is potentially missing variables influencing the baseline hazard. These are called random effects or frailties. Box-Steffensmeier and Jones (2004) show that omitting such covariates leads to the underestimation of the factors positively influencing the hazard rate and to an overestimation of the factors negatively relating to it. In our case, such factors are most likely related to specific locations and their relations to others. We account for these unobserved factors by means of including frailty terms into the model. These can either be individual or shared frailties. The first accounts for unit-specific effects and the second incorporates the effects relating to the clustering of observations, i.e. observations sharing the same region.

To do this, we make use of a Bayesian framework, which is particularly useful when applying survival models in spatial settings, i.e., when the observations are organized in a finite number of spatial units such as regions (Zhou and Hanson, 2018). This is the case here, as we assign all observations to the 402 German districts (NUTS3). Subsequently, we implement spatial dependencies through the inclusion of an intrinsic conditionally autoregressive (ICAR) element. The ICAR represents a spatial neighborhood matrix E, with an element e_{ij} being 1 if observations *i* and *j* are neighbors and 0 otherwise. Hereby, we consider the possibility of neighboring observations having similar risk propensities based on unobserved covariates (spatial autocorrelation) (Darmofal, 2009).

The quality of the models is compared by the Log Pseudo Marginal Likelihood (LPML) and the Deviance Information Criterion (DIC). A model is superior if its LPML value is larger and if its DIC value is smaller than another model.

5.4.3 Empirical variables

Supply-Push

We model local supply-push by the time it takes for the first wind turbine to be installed in region *i* with $FIRST\ TURBINE_{t \rightarrow e,\, i}$ [19] being our dependent variable. This event is explained with the following independent variables.

To measure the effect of existing producers on the likelihood of wind turbine deployment in region i, we create the variable $PRODUCER_{e\text{-}5}$. This counts the number of producers that exist within a radius of 25, 50, 75, or 100 kilometers from the first turbine in region *i*. We decided to work with a radius instead of administrative region borders (e.g., NUTS3) because wind turbines are often deployed at the borders of such regions (Broekel and Alfken, 2015) and we expect producers to have an effect on the deployment in neighboring regions as well. Moreover, we introduce a time lag of five years, i.e., we count the producers that entered the market at least five years before the observed deployment. This is justified because the decision process of installing and the installation of a wind turbine itself take around three to seven years (Kammer, 2011). Therefore, a time lag is needed when investigating whether wind turbine manufacturers create a technological niche and push the product into the regional market, allowing the manufacturer to plan, manufacture, and install the wind turbine.

We also include the number of existing wind turbines within a radius of twenty kilometers ($TURBINES_{i,e\text{-}5}$) with a time lag of five years. The variable summarizes all wind turbines deployed five years before the event *e*, i.e., it captures the installation of prior wind turbines in region i. Considering this variable allows us to model two effects in the growth phase: learning and aversion. On the one hand, the existence of wind turbines can signal learning effects, which in turn imply reduced planning and construction time for further wind turbines. On the other hand, if several wind turbines already exist, citizens might prefer to prevent the installation of further wind turbines, leading to longer planning times. The chosen distance is in line with the current literature (Broekel and Alfken, 2015).

[19] Start of the observation time is one year before the first event occurs.

Demand-Pull

With the local demand-pull model, we are interested in analyzing the effect of local demand on the propensity of wind turbine manufacturers emerging in region *i*. That is, we seek to explain the time it takes for a new company being founded in region *i*, which is captured by the dependent variable *FIRST PRODUCER*$_{t \rightarrow e,i}$.[20]

The first explanatory variable *TURBINES*$_e$ measures the number of wind turbines within a radius of 25, 50, 75, or 100 kilometers existing in the year of the founding of a wind turbine manufacturer. It approximates the demand conditions five years ago, which might have been the time at which the manufacturer made the decision to eventually found a firm. Planning and installing wind turbines, as well as establishing a business, is a long process usually requiring multiple years. Second, we approximate current demand conditions by counting the number of wind turbines installed in the five years following the emergence of a wind turbine manufacturer (*FUTURE TURBINES*$_{i,e+5}$). We make use of the same radii as with *TURBINES*$_e$. Due to the planning time of up to seven years, it seems likely that these turbines are publicly announced approximately five to six years before they start to operate. Hence, they are considered as current as well as short-term future demand that a manufacturer can satisfy.

Recent studies have shown that the emergence of new firms and industries greatly depends on spatial clustering and the existence of related variety (Boschma and Wenting 2007; Porter 1998). We construct the variable *PRODUCER*$_{i,e}$ that sums all wind turbine manufacturers already existing in region *i* when a new manufacturer emerges. A positive finding for this variable is in line with the idea of spatial clustering being important for a firm's emergence. The potential impact of related variety is capture by *RELATED*$_{i,e}$, which approximates the technological relatedness of the wind industry to other industries. It is based on the cosine similarity of the 4-digit IPC class F03D ("Wind motors") with all other IPC classes. More precisely, in a common manner, we estimate the cosine similarity based on the co-occurrence frequencies of 4-digit IPC classes on patents to obtain a measure of technological relatedness of each IPC class pair. In a second step, we calculate the revealed technological advantage (RTA) for each IPC class and region. If the RTA > 1 for IPC class F03D in region *i*, it reveals that this region patents more wind motors than the average region (Hidalgo et al., 2007). The two matrices, technological relatedness and RTA, are then multiplied with each other in order to obtain the aggregated relatedness coefficient for F03D and each region (see Neffke et al., 2011).

[20] Start of the observation time is one year before the first event occurs.

Control Variables

The diffusion of wind turbines is likely to be impacted by additional factors that are not in the focus of the present chapter. Crucially, the average wind speed in region *i* ($WIND_i$) is a natural candidate in this respect. The higher its level, the more electricity can be produced by the wind turbine and, hence, profitability increases making the installation more likely (Burton et al., 2011). The likelihood of installing wind turbines in regions is also determined by available space ($AREA_i$). Wind turbine installations need space to build up the necessary infrastructure, such as foundations and network access. Second, wind turbines have to be distant from other objects like buildings or trees, allowing for unhindered wind flow and avoiding externalities (Burton et al., 2011). We obtained data on the extent of residential areas and forests in regions from the German Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). We subtract from the total area of region *i* the areas already in use and the areas not available for wind energy due to natural constraints like rivers. As this variable is not measured on a yearly basis and the first figures are available for 1996, we decided to use its values from the last available year in each industry life-cycle phase, because we assume the available space decreases over time. On this basis, we estimate two values for the variables *AREA* for each NUTS3 region. The first represents these conditions in the initial phase of the industry (the year 2000)[21] and the second those at the end of the growth phase (the year 2010).

Moreover, we consider the political preferences toward sustainable energy in a region, which is crucial for the installation of wind turbines (Theyel, 2012). We approximate this with the share of votes for the Green party in federal elections ($GREEN_{i,e}$). Further regional factors are the population density ($POP_{i,e}$) and gross domestic product ($GDP_{i,e}$). In less populated regions it might be easier to plan and activate wind turbines as fewer people feel distracted by them. At the same time, a larger population represents more potential entrepreneurs that can establish wind turbine manufacturers. All three variables were gathered from the German database "GENESIS".

Moreover, for the demand-pull models, we control for regions being in the north of Germany ($NORTH_i$). More precisely, $NORTH_i$ is a dummy variable having the value of one if region *i* is part of the federal states of Hamburg, Bremen, Lower Saxony, Schleswig-Holstein, or Mecklenburg-Western Pomerania, and is zero otherwise. While we exclusively consider onshore wind turbines, the offshore business has become important for wind turbine

[21] There were no values for 1999.

manufacturers in recent years. Consequently, offering direct access to the sea is a beneficial attribute of North German regions (Fornahl et al., 2012). Finally, in the supply-push models, we consider the potential peculiarities of East Germany with dummy variable $EAST_i$ which is zero for all regions in West Germany and one for all East German regions. The variable controls for the East German regions becoming accessible to manufacturers at a later time and these regions being characterized by economic catching-up processes in the 1990s.

Time Dependency

Many of our covariates changed over the years. For example, *GREEN* changes every four years with the parliamentary elections. We therefore model time dependencies as described in Zhou et al. (2016) and Therneau et al. (2017). That is, we organize the data in time intervals that are defined in a way that variables' values change between intervals but not within. For the 1980s we lack data for several variables (see Table 5.1) and therefore define variables' values based on the earliest available observation.

Table 5.1: Variable operationalization and source

	Variable	Description	Operationalization	Source	Year
Dependent	FIRST TURBINE$_{t \rightarrow e, i}$	Wind turbine	Time of first wind turbine installed in region i (NUTS 3).	Energy-map.org	Year of event (t)
	FIRST PRODUCER$_{t \rightarrow e, i}$	Wind farm producer	Time of first producer entering the market in region i (NUTS 3).	Various websites	Year of event (t)
Independent	TURBINES$_{e}$	Existing wind turbines	Existing wind turbines within a radius of 20 km at time e	Energy-map.org	Year of event (t)
	TURBINES$_{e-5}$		Existing wind turbines within a radius of 25, 50, 75 or 100 km at time e-5	Energy-map.org	< t-5
	FUTURE TURBINES$_{i,e+5}$	Future wind turbines	Wind turbines within a radius of 25, 50, 75 or 100 km at time e+5	Energy-map.org	t+1 to t+5
	DISTANCE$_{P,FT}$	Distance in km	Distance between new producer and future wind turbines		
	PRODUCER$_{e-5}$	Existing wind farm manufacturers	Wind farm producers within a radius of 25, 50, 75 or 100 km in time e-5	Various websites	< t-5
	PRODUCER$_{i,e}$		Wind farm producers in region i (NUTS 3) at time e	Various websites	Year of event (t)
	DISTANCE$_{T,P}$	Distance in km	Distance between new turbines and existing producers		
	WIND$_{i}$	Average wind speed	Above average wind speed (ger.: Windhöffigkeit) in region i.	German weather service	1981 - 2000
	AREA$_{i}$	Available space	Available space in region *i* useable for wind turbines	Regional database Germany	1998; 2010
	GREEN$_{i,e}$	Votes	Above average percentage of votes for Bündnis 90/The Greens in federal elections in region i at time e	Regional database Germany	1994; 1998; 2002; 2005; 2009
	GDP$_{i,e}$	Regional GDP	Above average gross domestic product in region i at time e	Regional database Germany	1992 – 2010
	POP$_{i,e}$	Population density	Above average population density of region i at time e	Regional database Germany	1992 – 2010
	RELATED$_{i,e}$	Patent relatedness of region i	Co-occurrence of IPC classes with class F03D ("Wind motors") in a patent and location quotient for region i at time e	PATNET	1974 - 2010
	NORTH$_{i}$	Regions of North Germany	Dummy variable: 1 if region i belongs to Hamburg, Bremen, Lower Saxony, Schleswig-Holstein or Mecklenburg-Western Pomerania, otherwise 0		
	EAST$_{i}$	Region in East Germany	Dummy variable: 1 if region i belongs to Mecklenburg-Vorpommern, Brandenburg, Saxony, Saxony-Anhalt and Thuringia, otherwise 0		

5.5 Results

Before coming to the actual results, please note that we do not report p-values but rather the 95% confidence intervals, which are more common when using Bayesian event models (Darmofal, 2009; Craioveanu and Terrell, 2016). Moreover, the variable capturing spatial dependencies (ICAR) is significant in all estimations, justifying the choice of our empirical approach.

5.5.1 Initial phase

Local supply

The results for the initial phase are presented in Table 5.2. The models explain the effect of local supply on the creation or activation of local demand (supply-push), i.e., the likelihood of wind turbine installations.

$WIND_i$ is significant and positive, implying that regions with high wind speeds are likely to be among the first to install wind turbines. If wind levels increase by one meter per second, the probability of a first wind turbine being installed increases by 124%. This finding supports the idea of turbines being expensive and having low degrees of efficiency in the industry's initial phase (Neukirch, 2010; Kammer, 2011). This made early installments more attractive in regions with high wind speeds.

$POP_{i,e}$ is also significant and positive. This contradicts our expectations that wind turbines are more likely to be installed in less populated regions to reduce land-use conflicts (see e.g., Short (2002) and the "nimbyism" discussion). However, as our models explain the timing and not the number of wind turbines, we see this finding to be in line with the technology diffusion literature. This literature argues that new technologies are more likely to emerge in urban regions and that their diffusion starts from more central places (Hägerstrand, 1952).

$\mathrm{GDP}_{i,e}$ is significant and negative, i.e. regions with large gross domestic products per capita are less likely to be early locations for wind turbine installations. This fits with the northern German regions, which were found to be more likely locations for wind turbines.

In the model considering a radius of one hundred kilometers, $PRODUCER_{e-5}$ is significant and negative. This suggests that the presence of wind turbine producers tends to increase the time needed to install the first wind turbine in a region. This result is surprising because we observe a relatively strong overlap between producers' and wind turbines' locations in North Germany (see Figure 5.5). However, again, our models capture the timing of turbine installations and their locations. Accordingly, this finding suggests that the earliest wind turbines' locations were not proximate to those of producers, which is supported by the fact that the average distance between producers and wind turbines is 81 km. Hence, while they share the general location of northern Germany, early turbines were not installed directly at the producers' locations but in more suitable places in the wider surroundings.

Table 5.2: The regional supply model in the initial phase

	Local Supply-Push Initial Phase (1983 - 1998) 50 km radius	100 km radius
Supply-Push		
$PRODUCER_{e-5}$	0.0079	**-0.359**
	(-0.184, 0.194)	**(-0.503, -0.211)**
$DISTANCE_{T,P}$	-	-
$TURBINES_{e}$	-	-
Regional Characteristics		
$WIND_{i,e}$	**0.492**	**0.808**
	(0.337, 0.675)	**(0.552, 1.039)**
$AREA_{i,e}$	0.003	0.045
	(-0.011, 0.07)	(-0.031, 0.114)
$GREEN_{i,}$	-5.743	-4.517
	(-1.42, 2.371)	(-15.4, 5.097)
$GDP_{i,e}$	**-0.008**	**-0.01**
	(0.0003, 0.0031)	**(-0.015, -0.0043)**
$POP_{i,e}$	**0.0017**	**0.0025**
	(0.0003, 0.003)	**(0.0002, 0.005)**
EAST	-0.384	**-1.777**
	(-0.025, 0.387)	**(-2.656, -0.683)**
ICAR	**1.641**	**10.79**
	(0.718, 2.975)	**(5.71, 18.00)**
Survival Model	Proportional hazards	Proportional hazards
LPML	-527	-432
DIC	1051	812
N | Events	2159 | 152	2159 | 152

Cell entries are the posterior means, with 95% credible intervals in parentheses

To test the robustness of our results, we calculated additional models with an alternative time lag of three years and alternative radii of 25 and 75 km. The results are reported in the first two rows in Table A5.1 in the appendix. Our results are robust with respect to the specification of the time lag; there are no visible differences when using three or five years. By and large, these models also confirm our main findings with $PRODUCERS_{e-5}$ being negative and significant. $PRODUCERS_{e-5}$ remains insignificant in the models in which we consider wind turbine installations within a radius of 25 km and 50 km to district centers. However, this is explained by the fact that few wind turbines exist within a 50 km distance of producers, as the average minimum distance between wind turbines and producers is 81 km (Table A5.2). We also tested an alternative specification considering the distance between the first wind turbine and the closest existing producer. Unfortunately, these models did not converge. For this reason, we also calculated a Cox proportional hazard model with the same frailties (see Table A5.3).

Local demand

We now turn toward the question of whether local demand for wind turbines attracts the emergence of producers. The results of our estimations are shown in Table 5.3.

RELATED$_i$ has a significantly positive coefficient. Regions with a related patent portfolio have a higher likelihood of wind turbine producers emerging. This confirms existing works on the impact of relatedness of regional diversification and firm foundation (Neffke et al., 2011). It also fits with Kammer (2011) who highlights that the German wind turbine producers benefited from the shipbuilding industry and with Breul et al. (2015) who also find a positive relationship between technological relatedness and the time of local manufacturers' emergence. Accordingly, this chapter adds further support for the importance of relatedness for industrial development.

Table 5.3: Demand-pull results of the initial phase

	Local Demand-Pull Initial Phase (1974 - 1998)		
	50 km	**100 km**	**Distance**
Demand-Pull			
FUTURE TURBINES$_{e+5}$	**0.0053 (0.0008, 0.009)**	0.0009 (-0.0004, 0.003)	-
DISTANCE$_{P,FT}$			-0.003 (-0.008, 0.0013)
TURBINES$_e$	-0.032 (-0.086, 0.012)	-0.017 (-0.0493, 0.024)	-
Related Variety			
PRODUCER$_{i,e}$	-	-	-
RELATED	**0.014 (0.006, 0.002)**	**0.0014 (0.0005, 0.002)**	**0.0014 (0.0006, 0.003)**
Regional Characteristics			
NORTH$_i$	**1.329 (0.329, 2.205)**	**1.353 (0.177, 2.454)**	**1.329 (0.177, 2.454)**
GREEN$_{i,e}$	2.970 (-1.288, 1.730)	2.308 (-0.162, 0.179)	2.308 (-16.29, 0.177)
GDP$_{i,e}$	-0.003 (0.008, 0.002)	-0.004 (-0.01, 0.0016)	-0.0042, (-0.0101, 0.0016)
POP$_{i,e}$	0.0002 (-0.002, 0.002)	0.0003 (-0.002, 0.003)	0.0003 (-0.0018, 0.0022)
ICAR	**0.775 (0.05, 2.97)**	**2.919 (0.042, 12.855)**	**0.676 (0.06, 2.591)**
Survival Model	Proportional Hazards	Proportional Hazards	Proportional Hazards
LPML	-235	-222	-236
DIC	468	441	470
N | Events	7288 | 57	7288 | 57	7288 | 57

Cell entries are the posterior means, with 95% credible intervals in parentheses

Furthermore, we find more favorable conditions for this industry's development in northern regions, which we captured with the variable $NORTH_i$. This variable is significantly positive and indicates that the likelihood of the first manufacturer entering a region increases by a factor of three when the region is located in the north of Germany. We interpret this as an indication of the relevance of access to offshore activities for wind turbine manufacturers' locations and for the generally more wind turbine-friendly conditions in northern Germany (a more suitable landscape and more sparsely populated) that are not captured by the other variables in our model.

In contrast to PV manufacturers that appear to prefer urban regions as locations (Breul et al., 2015), population density remains insignificant in our investigations and thus appears to be irrelevant for wind turbine manufacturers. With all other factors equal, new facilities have equal chances of being located in urban and rural regions. Consequently, these regions do not appear to offer any particular locational benefits or disadvantages. The latter is somewhat surprising, as the transportation of towers and rotor blades would seem to be more cumbersome and expensive in urban environments and hence, make these locations less attractive.

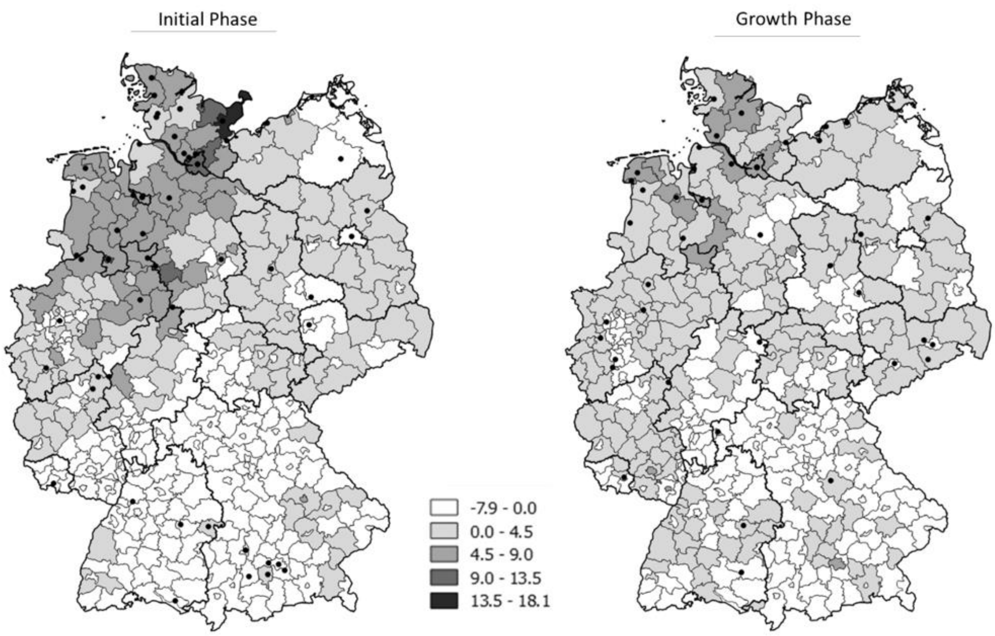

Figure 5.5: Spatial distribution of wind turbines in Germany (dark grey: high propensity of installation, white: low propensity, black points: producer)

With respect to the relevance of a local demand-pull effect, we look at the coefficients of *FUTURE* $TURBINES_{e+5}$, $DISTANCE_{P,FT}$, and $TURBINES_e$ that approximate a local demand-pull effect. *FUTURE* $TURBINES_{e+5}$ is significantly positive in the 50 km model with the coefficients of $DISTANCE_{P,FT}$ and $TURBINES_{i,e}$ remaining insignificant. The emergence of a producer is 0.5% more likely in regions that see one more additional wind turbine being installed over the next five years. The finding provides some support for the idea that local demand stimulates the emergence of wind turbine producers. Note that the effect appears smaller than in reality because wind turbines are often installed in larger numbers at the same time, which multiplies this effect.

5.5.2 Growth phase

Local supply

During the industry's growth phase, the installation of wind turbines was impacted by favorable natural conditions, which is shown by the significantly positive coefficients of wind speed $WIND_i$ and available space $AREA_i$ (see Table 5.4). Higher wind speeds and more available space decreases the time until the installation of a first wind turbine. Over time, the relevance of $AREA_i$ increases as available open spaces become more scarce and therefore potentials of utilization conflicts grow (Burton et al., 2011). As in the initial phase, $GDP_{i,e}$ is significantly negative: wind turbines are installed earlier in economically weaker regions.

Interestingly, in the model considering a 50km radius to wind turbines, $GREEN_{i,e}$ is significant but negative, that is, regions with larger shares of green voters needed more time to install their first wind turbine. This seems contradictory at a first glance as wind power is a renewable, "green" energy source. However, voters favoring ecological behavior and thus voters of green parties are more frequently found in South German cities, such as Freiburg, Stuttgart, or Munich. They are significantly underrepresented in the northern regions where most wind turbine installations took place. In addition to a strong south-north discrepancy, voting for the Green party is also much more frequent in the largest cities. Accordingly, this variable may also capture that wind turbine installations are much less common there. Moreover, wind turbine installations also represent interference with local ecological systems, which these types of voters likely oppose. As the variable captures multiple aspects, its relatively large coefficient is not surprising (coefficient: -7.72). Given that *GREEN's* value ranges between 0.02 and 0.29, an increase of 1% in votes decreases the likelihood of a first wind turbine installation in a region by about 7%.

$PRODUCER_{e-5}$, which approximates local supply-push processes, is insignificant in both models. However, the distance to the next producer, $DISTANCE_{T,P}$ is significantly positive. Hence, it takes more time for the first wind turbine to be installed in a region proximate to an existing manufacturer. This clearly does not correspond to a significant local supply-push effect in this phase. Rather, the contrary appears to hold, the first wind turbines are installed at greater distances to manufacturers. However, this result could be driven by some producers being located far away from wind turbines' early locations, i.e. in the south of Germany.

Table 5.4: Supply-push results for growth phase

	Local Supply-Push Growth Phase (1999 - 2010)		
	50 km	**100 km**	**Distance**
Supply-Push			
$PRODUCER_{e-5}$	0.198 (-0.493, 0.001)	-0.374 (-0.124, 0.068)	
$DISTANCE_{T,P}$			**0.033** **(0.0195, 0.0401)**
$TURBINES_e$	-0.017 (-0.043, 0.008)	-0.012 (-0.026, 0.001)	-0.019 (-0.409, 1.361)
Regional Characteristics			
$WIND_{i,e}$	**0.261** **(0.083, 0.531)**	**0.247** **(0.051, 0.384)**	**0.40** **(0.212, 0.739)**
$AREA_{i,e}$	**0.333** **(0.201, 0.453)**	**0.252** **(0.190, 0.335)**	**0.434** **(0.351, 0.544)**
$GREEN_{i,}$	**-7.72** **(-14.24, -0.445)**	-6.169 (-0.139, 0.658)	-6.151 (-19.51, 2.831)
$GDP_{i,e}$	**-0.009** **(-0.013, -0.006)**	**-0.009** **(-0.013, -0.004)**	**-0.015** **(-0.0202, -0.0103)**
$POP_{i,e}$	0.0015 (-0.0003, 0.003)	0.0013 (-0.0006, 0.006)	**0.0032** **(0.0013, 0.0049)**
EAST	-0.014 (-1.466, 1.162)	0.487 (-0.347, 1.757)	0.378 (-0.434, 1.215)
ICAR	**23.084** **(3.724, 39.045)**	**7.804** **(1.257, 16.046)**	**53.24** **(39.941, 69.899)**
Survival Model	Proportional hazards	Proportional hazards	Proportional hazards
LPML	-510	-572	-400
DIC	906	1101	655
N | Events	1940 | 204	1940 | 204	1940 | 204

Cell entries are the posterior means, with 95% credible intervals in parentheses

Therefore, as a robustness test, we included only regions of North Germany. The coefficient of $DISTANCE_{T,P}$ becomes insignificant (Table A5.4 in the appendix). In any case, the finding clearly suggests that better possibilities to build local support for wind turbines or lower transportation costs to potential wind turbine locations did not matter at this stage of the industry's life-cycle since more suitable natural conditions for wind turbine installations were more important.

Local demand

The results for the growth phase are also reported Table 5.5. The number of existing producers in a region *($PRODUCERS_e$)* is significantly positive, which indicates that the industry is concentrating in space as existing producers attract further producers. This may either be due to spin-off processes (Klepper, 2006) or because of the increasing relevance of Marshallian externalities (Neffke et al., 2011). As $RELATED_i$ is insignificant it implies the irrelevance of related industries at this stage. This phase of the industry's life-cycle is also characterized by firms opening up second production locations with little to no R&D activities (Kammer, 2011). For these, the presence of related knowledge is less relevant than advantages linked to shared infrastructure and labor pooling effects.

Table 5.5: Demand-pull results of the growth phase

	Local Demand-Pull Growth Phase (1999 - 2010)		
	50 km	**100 km**	**Distance**
Demand-Pull			
FUTURE $TURBINES_{e+5}$	0.001 (-0.004, 0.005)	-0.003 (-0.002, 0.0009)	
$DISTANCE_{P,FT}$	-	-	**-0.073 (-0.144, -0.009)**
$TURBINES_e$	0.0001 (-0.016, 0.015)	0.005 (-0.007, 0.018)	0.0009 (-0.013, 0.013)
Related Variety			
$PRODUCER_{i,e}$	**0.692 (0.186, 1.244)**	**0.601 (0.239, 0.951)**	**0.610 (0.242, 0.941)**
RELATED	0.0006 (-0.002, 0.003)	0.0004 (-0.002, 0.002)	0.0005 (-0.001, 0.002)
Regional Characteristics			
NORTH	1.084 (-0.114, 2.227)	**1.275 (0.418, 2.142)**	**0.987 (0.166, 1.805)**
$GREEN_{i,e}$	-7.417 (-22.53, 5.228)	-8.637 (-20.69, 3.630)	-7.580 (-20.78, 3.497)
$GDP_{i,e}$	0.0008 (-0.003, 0.0044)	0.0002 (-0.004, 0.004)	0.0007 (-0.003, 0.004)
$POP_{i,e}$	**0.0023 (0.0002, 0.004)**	0.002 (-0.0003, 0.003)	0.0018 (-0.00009, 0.0037)
ICAR	**1.608 (0.1347, 5.7041)**	**0.067 (0.005, 0.195)**	**0.080 (0.006, 0.464)**
Survival Model	Proportional Hazards	Proportional Hazards	Proportional Hazards
LPML	-192	-191	-188
DIC	378	381	375
N | Events	4648 | 46	4648 | 46	4648 | 46

Cell entries are the posterior means, with 95% credible intervals in parentheses

$POP_{i,e}$ is significant and positive in one model (50 km radius) meaning that regions with higher populations are more likely to be the first to witness an early firm entry. This result fits with the urbanized regions having larger potentials of entrepreneurial activities (Boschma and Wenting, 2007).

$TURBINES_{i,e}$ and $FUTURE\ TURBINES_{i,e+5}$ remain insignificant in these models. The number of existing wind turbines and those that will be built in the next five years within a certain distance to the producer does not explain producers' emergence. However, we do find a significantly negative coefficient of $DISTANCE_{T,P}$, which supports the demand-pull hypothesis. During the growth phase, new producers are more likely to emerge early in regions geographically proximate to future wind turbines. With every additional kilometer a region is more distant to future wind turbines, the probability of a producer emerging decreases by about 7%. This is likely because firms seek to minimize the transport costs of their final products. In fact, Klepper (2006) presents evidence of firms in the automobile and tire industries locating new production facilities in geographic proximity to customers to save transportation costs. In the case of wind turbines, these costs are even higher and therefore production facilities located close to demand are more valuable. However, it may also be the case that firms will locate near demand to better interact with their customers and understand their preferences (Fabrizio and Thomas, 2012). Crucially, while there are benefits to geographic proximity, manufacturers do not need to be immediately co-located with wind turbine locations, as the insignificance of $TURBINES_{i,e}$ and $FUTURE\ TURBINES_{i,e+5}$ underlines. This finding is reasonable given that wind turbines require open space and a certain distance to settlements.

5.6 Discussion and conclusion

The aim of this chapter was to analyze whether and to what extent local supply-push or local demand-pull mechanisms characterize an industry's spatial evolution. We have combined supply-side arguments from the field of economic geography (e.g., Weber, 1909; Myrdal, 1957; Frenken et al., 2007; Boschma and Frenken, 2011) with more demand-oriented arguments from the literature on technological systems (Geels, 2004; Geels and Schot, 2007). In particular, we argued that the relevance of local supply-push and demand-pull factors changes over the life-cycle of an industry.

These arguments were tested using the example of the German wind turbine industry and its spatial evolution over its initial and growth phases. Empirically, we made use of data on manufacturers' founding dates and their locations. This was merged with information on the installation time and geographical locations of wind turbines. We employed Bayesian survival models to test the relative importance of the local supply-push and local demand-pull hypotheses.

The local supply-push models, which analyze the spatial diffusion of wind energy turbines, highlighted the importance of regions' natural conditions. In order to maximize turbines' revenues, the first turbines are installed in regions with high wind speeds and with available open spaces. We find these natural conditions to be more important for the success and diffusion of this technology than manufacturers establishing local supportive niches in the initial and growth phase of the industry. Accordingly, local supply-push induced by manufacturers appears to be less relevant for this industry's evolution. However, this is not to say that these processes have not been present at all, for instance, some regions can be seen as testing areas for early wind turbine installations (Kammer, 2011). According to our results, this does not seem to have been essential for the evolution of the industry spatially.

In the local demand-pull models that analyze the spatial evolution of wind turbine manufacturers, we observe the expected processes typical for the industrial phases of emergence and concentration. In the initial phase of the industry, manufacturers are more likely to emerge in regions with related knowledge, for example, in which the shipbuilding industry was present. Over time, relatedness becomes less important and new firms tend to emerge in geographic vicinity to already existing manufacturers, which fosters the industry's spatial concentration. In addition, local demand-pull becomes more relevant: with every additional kilometer to demand (future wind turbine installations), a new producer is less likely to emerge in a region. In summary, we confirm the importance of related variety, urbanization, and industrial agglomeration shaping the industry's spatial distribution. In addition to providing further evidence for this, this chapter highlights the significant role of local demand.

These characteristics of the wind industry are in line with the argumentation of Binz and Truffer (2017) who define four kinds of global innovation systems (GIS). In particular, they classify the wind industry as a "spatially sticky GIS". Accordingly, the wind industry is built upon specialized user needs and experience-based skills that are hard to copy and that are unlikely to diffuse in space. Other examples of such industries are biogas, luxury watchmaking, or legal services (Ibid.). Crucially, in the growth phase, these industries' spatial distributions remain similar to those of the initial stage. With the exception of East Germany expanding the set of potential locations in 1989, the results support this view, as the spatial distribution of producers remained similar over time and as the existence of local producers was identified to contribute to the establishment of new producers. In contrast, the photovoltaic industry, which was subject to similar politically support as the wind industry (Jacobsson and Lauber, 2006; Dewald and Truffer, 2012; Breul et al., 2015), is characterized as a "footloose GIS" (Binz and Truffer, 2017). This implies that regions diversifying into the PV industry at an early stage of this industry's life-cycle are much more in danger of losing their initial market dominance. In fact, this is what happened in Germany, where regions with an initial advantage in this industry were overtaken (and almost completely eliminated from the market) by other locations in Asia that entered the industry at a later stage. Consequently, any generalization of our results is conditional on such differences between industries and the respective GIS.

Our empirical study has several shortcomings, which may lead the way for future research. First, the German wind industry has specific and partly unique characteristics; it is an industry highly supported by the government implying that its evolution is strongly related to decisions made in politics (Johnson and Jacobsson, 2003; Kaldellis and Zafirakis, 2011). We are confident that this did not significantly impact producers' and wind turbines' precise locations, as nationwide tax and tariff schemes have been mostly used as policy tools in Germany. These shaped demand and supply in general but not their spatial distributions. Future research should nevertheless reevaluate the arguments using industries more independent of policy influence. Second, to generate a quantitative empirical setting, we focused on the push and pull mechanisms of technology. However, the literature emphasizes technology transitions entailing an interplay of societal rules, norms, market preferences, and policy (see for example Schot and Geels, 2008). Moreover, Fabrizio and Thomas (2012) discovered that in the pharmaceutical industry, access to local demand changes firms' patterns of innovation as it eases understanding local peculiarities in customer needs. Hence, looking in more detail into these processes is likely to reveal further interesting procedures that remain hidden in our research design. Third, we did not explicitly consider the main product of the industries (wind turbines) to be heterogeneous and evolving over time as well. While certain features of wind turbines can be regarded as a dominant design (3-blade rotor), there are substantial technological variations (e.g., gear vs. gearless designs). Treating all wind turbines to be alike therefore represents a significant assumption on which our empirical study is built. Fourth, we lack the information of which wind turbines were manufactured by which producers. With such information at hand, it would be possible to test if producers only create local demand for their own products or whether they also open markets for competitors**.** Lastly, our quantification of local demand by means of wind turbines installed in regions in subsequent periods does not directly reflect the demand curve. Rather it may represent the point of supply and demand levels being equal. Accordingly, we might underestimate actual demand in regions in which it exceeds supply capacities. However, we argue and provide evidence for supply being spatially mobile to some extent. This is particularly true when looking at longer time periods. In these cases, manufacturers may emerge or re-locate in proximity to demand which implies that the equality of supply and demand levels in subsequent periods may also reflect the "excess" of demand in previous periods. How to properly approximate and quantify demand at the local level needs to be addressed by future research in more detail. With improved access to novel and more precise data in the future, we are confident that the demand side will receive more attention in studies investigating the evolution of industries in time and space.

Accordingly, while the present chapter represents an additional step in disentangling the interplay between demand and supply in industry’s evolution in time and space, we are still far away from fully understanding these processes.

Appendix

5.A1 Robustness checks

Our main variable PRODUCERS stays always negative and significant, however, in some models it is insignificant (Table A5.1). This is true for the 25 km and 50 km radii. When taking a look at Table 7, we see that the average minimum distance between an installed turbine and producer is at least 57 km. Accordingly, the number of producers within 25 km or 50 km is generally very low, which is the most likely reason for its frequent insignificance.

There is no difference in results with respect to the time lags of three and five years.

Table A5.1: Main results of the robustness test

Model	Phase		Producers (25 km)	Producers (50 km)	Producers (75 km)	Producers (100 km)	Distance
Supply-Push	Initial	3 years and earlier	Insignificant	Insignificant	**Negative Significant**	**Negative Significant**	Not converged
		5 years and earlier	Insignificant	Insignificant	**Negative Significant**	**Negative Significant**	Not converged
Supply-Push	Growth	3 years and earlier	Insignificant	Insignificant	**Negative Significant**	Insignificant	**Positive Significant**
		5 years and earlier	Insignificant	Insignificant	**Negative Significant**	Insignificant	**Positive Significant**

Model	Phase		Future Turbines (25 km)	Future Turbines (50 km)	Future Turbines (75 km)	Future Turbines (100 km)	Distance
Demand-Pull	Initial	Within 3 years	Insignificant	Insignificant	Insignificant	Insignificant	Insignificant
		Within 5 years	**Positive Significant**	**Positive Significant**	Insignificant	Insignificant	Insignificant
Demand-Pull	Growth	Within 3 years	Insignificant	Insignificant	Insignificant	Insignificant	Insignificant
		Within 5 years	Insignificant	Insignificant	Insignificant	Insignificant	**Negative Significant**

The demand-pull models show different results when it comes to three- or five-year time lag and our main variable FUTURE TURBINES. The three-year model of the initial phase has only insignificant values. We trace this result back to the fact that only very few turbines are deployed within three years (see Table A5.2). Again, there is no difference in the direction of significance.

Table A5.2: Average Values of turbines and producers within a given radius (median values in brackets)

Model	Phase		Producers 25 km	Producers 50 km	Producers 75 km	Producers 100 km	Distance
Supply-Push	Initial	3 years and earlier	0.16 (0)	0.69 (0)	1.49 (1)	2.48 (2)	77.39 km (62.54 km)
		5 years and earlier	0.15 (0)	0.62 (0)	1.35 (1)	2,23 (2)	81.72 km (68.41 km)
Supply-Push	Growth	3 years and earlier	0.31 (0)	1.11 (1)	2.41 (2)	3.79 (3)	57.41 km (48.83 km)
		5 years and earlier	0.27 (0)	0.98 (1)	2.158 (2)	3.42 (3)	62,46 km (52.73 km)

Model	Phase		Future Turbines 25 km	Future Turbines 50 km	Future Turbines 75 km	Future Turbines 100 km	Distance
Demand-Pull	Initial	Within 3 years	4.45 (0)	17.45 (0)	39.31 (2)	68.98 (3)	131,23 km (51.81 km)
		Within 5 years	34.54 (0)	77.63 (2)	136.00 (6)	104.63 (11)	104,62 km (31.33 km)
Demand-Pull	Growth	Within 3 years	20.17 (10)	79.7 (50)	178.4 (122)	311.2 (227)	13.97 km (10.92 km)
		Within 5 years	28.82 (15)	114.30 (78)	256.00 (194)	446.00 (359)	11.45 km (9.17 km)

Table A5.3: Results of the Cox proportional hazard models

	Local Supply-Push Initial Phase (1983 - 1998)			
	Distance (3-years lag)		Distance (5-years lag)	
	Coefficient (SE)	p-value	Coefficient (SE)	p-value
Supply-Push				
$DISTANCE_{T,P, e-3}$	-0.002 (0.002)	0.403		
$DISTANCE_{T,P, e-5}$			-0.0012 (0.002)	0.525
Regional Characteristics				
$WIND_{i,e}$	0.504*** (0.063)	< 1e-3	0.505*** (0.064)	< 1e-3
$AREA_{i,e}$	0.058*** (0.019)	0.003	0.059*** (0.019)	0.002
$GREEN_{i,}$	-10.78*** (3.864)	0.005	-10.76e*** (3.855)	0.005
$GDP_{i,e}$	-0.007*** (0.002)	< 1e-3	-0.007*** (0.002)	< 1e-3
$POP_{i,e}$	0.001*** (0.001)	0.014	0.001** (0.001)	0.016
EAST	2.226*** (0.318)	< 1e-3	2.218*** (0.318)	< 1e-3
ICAR	-0.011 (0.001)	< 1e-3	-0.011*** (0.987)	< 1e-3
Concordance	0.821		0.821	
Likelihood ratio test	257 on 8 df		212 on 8 df	

Table A5.4: Robustness check of distance. Data only includes regions of North Germany (i.e. regions of the states Hamburg, Bremen, Lower Saxony, Schleswig-Holstein or Mecklenburg-Western Pomerania)

	Local Supply-Push Growth Phase (1999 - 2010) Distance (3-years lag)	**Distance (5-years lag)**
Supply-Push		
$DISTANCE_{T,P, e-3}$	-0.013 (-0.035, 0.009)	
$DISTANCE_{T,P, e-5}$		-0.016 (-0.039, 0.016)
$TURBINES_{e}$	0.006 (-0.007, 0.019)	0.006 (-0.016, 0.031)
Regional Characteristics		
$WIND_{i,e}$	0.243 (-0.123, 0.605)	0.281 (-0.186, 0.624)
$AREA_{i,e}$	**0.152** **(0.071, 0.237)**	**0.165** **(0.064, 0.352)**
$GREEN_{i,}$	-5.545 (-1.783, 7.790)	-7.854 (-24.45, 8.782)
$GDP_{i,e}$	**-0.005** **(-0.009, -0.0002)**	**-0.006** **(-0.011, -0.001)**
$POP_{i,e}$	0.0001 (-0.002, 0.002)	0.001 (-0.002, 0.003)
EAST	0.865 (-1.286, 3.006)	0.831 (-1.402, 3.506)
ICAR	**3.503** **(0.9452, 8.4271)**	**3.61** **(0.5127, 12.6271)**
Survival Model	Proportional Hazards	Proportional Hazards
LPML	-42	-60
DIC	61	104
N | Events	169 | 68	169 | 68

Conclusion

Researchers in economic geography strive to understand the changes in economic landscapes and the reasons for regional divergence. Therefore, they explore the evolution of industries and the generation of innovations as sources of new economic welfare and regional prosperity. In recent years, geographers have devoted much more attention to the initial creation of new technologies and products than to their subsequent diffusion. In many cases, however, it is not the generation of a new technology within a single region that changes the structure of economic landscapes; rather, it is its widespread diffusion and adoption (Grübler, 1997). Consequently, this thesis has focused on three dimensions of knowledge diffusion—technology, networks and regional context—to better understand the mechanisms of spatial diffusion processes. Related to these dimensions, four gaps in the literature have been recognized that have been investigated in the four central chapters of this work. Chapter 6 highlights the thesis' contributions, summarizes the empirical findings, and discusses the limitations as well as avenues for future research. Last but not least, it will derive policy implications.

6.1 Theoretical contributions

Combining the concepts of proximity, technological complexity and spatial patterns

To analyze the spatial diffusion of complex technologies, Chapter 2 has combined three theoretical concepts: proximities (Boschma, 2005), complexity (Simon, 1962; Kaufman, 1993) and spatial diffusion patterns (Hägerstrand, 1952, 1967). So far, these three concepts have not been considered together, although all three concepts are often used to study technology diffusion. Feldman et al. (2015) study the diffusion of biotechnology and therefore consider proximities and spatial diffusion patterns, while Balland and Rigby (2017) analyze the effects of geographical and technological proximity on the diffusion of complex technologies. Taking all three concepts into account contributes to our understanding of spatial technology diffusion.

First, including the perspective on spatial diffusion patterns allows an increase in understanding about geographic proximity beyond the common distinction of whether it influences diffusion processes or not. Therefore, hierarchical and contagious diffusion patterns have been discussed in Chapter 2. Besides these two patterns that have been introduced by Hägerstrand (1952, 1967), a further diffusion pattern has been added: leap-like diffusion. This pattern is characterized by a diffusion that leaps from one region to another without showing any kind of neighborhood effect.

Additionally, discussing the dimension of technological complexity promises to offer insight into why certain technologies diffuse hierarchically while others diffuse contagiously or in a leap-like form. Complex technologies tend to consist of more subcomponents, and more information is required to understand their structures than those of simple ones (Dehmer et al., 2009). These characteristics imply that simple and complex technologies diffuse with different spatial patterns. Simple technologies have fewer requirements for adoption and might therefore diffuse with far less friction. Geographic proximity is likely to play a minor role, leading to a leap-like diffusion pattern. Interestingly, high levels of technological complexity do not suggest one distinct spatial pattern per se. On one hand, the infrastructure and capability requirements of complex technologies indicate a hierarchical diffusion, as those requirements are more likely to be found in cities (Hagget, 2001; Bettencourt et al., 2007). On the other hand, the higher efforts necessary to understand complex technologies indicate the necessity of strong personal interactions that are more likely to form in geographic proximity (Boschma, 2005) and imply a contagious diffusion. Thereby, Chapter 2 has shown a way of connecting the traditional innovation diffusion literature (Hägerstrand, 1966; Blaut, 1977) with recent concepts of EEG.

The dissolution of network links

Although the structure and evolution of social networks has been extensively studied, the mechanisms underlying the dissolution of links has remained relatively unexplored. Some authors indicate that the dissolution is not just a strictly inverse process of link formation but is independent and, therefore, worthwhile to explore (Balland, 2012; Krivitsky and Handcock, 2014). However, a comprehensive theoretical discussion about the mechanisms of link dissolution is still lacking. Chapter 4 has presented a starting point for this discussion and has elaborated on dissolution processes.

The discussion considers location factors and aspects of the dyad as well as the structural level of networks. For example, it is argued that organizations in urban regions will likely dissolve links earlier to exploit the greater possibilities of partners that cities offer. Moreover, they will do this with an even higher probability if partners have high cognitive proximities. Cognitive proximity may enhance the formation of links because it enables the partners to exchange knowledge more efficiently (Nooteboom et al., 2007). But as cognitively proximate organizations could also be competitors (Boschma, 2005), cognitive proximity might simultaneously facilitate the dissolution of links and not its maintenance. In this sense, Chapter

4 contributes an important first step to the literature discussing the process of link dissolution as an elementary aspect of network evolution.

The demand side of industry emergence

The literature of innovation diffusion highlights the relevance of the regional context to adoption probabilities and diffusion processes (Blaut, 1977; Ormrod, 1990; Rogers, 2003). Regional characteristics shape the perception of potential adopters regarding the usability of new technologies and products. Thus, the regional context tends to form the demand of economic agents. Adding to this, technology transition studies emphasize the relationship of local demand and the creation of technological niches nurturing industries in their early phases (Geels, 2004). In economic geography, the processes and patterns of industry emergence in regions are frequently understood as the outcome of previously existing regional industry structure (Boschma and Martin, 2010). Thereby, the emphasis rests on the relatedness or unrelatedness of novel industries to current ones. Stated differently, the literature looks at the local supply of resources and infrastructure and their impact on attracting new industries (Boschma, 2017). Far less attention has been directed to the effect of local demand on industry dispersion. For this reason, Chapter 5 contributes an extensive elaboration on how demand shapes the location decisions of manufacturers.

The chapter has combined arguments from the literature of technological transition and innovation management with aspects of EEG. Thereby, Chapter 5 has distinguished between local supply-push and local demand-pull mechanisms. The first has been described as the creation of regional niches in which technologies are supported and nurtured (Schot and Geels, 2008). These niches are incubated by manufacturers to create small local markets in which their product can be introduced without facing complete competition (Jacobsson and Johnson, 2000). Those niches tend to be place specific as local constellations of actors and institutions shape these (Coenen et al., 2012). In contrast, manufacturers might not actively create local markets but may rather follow existing consumer demand. Hence, manufacturers develop product features according to customer preferences. In order to understand these best, close relationships to customers are necessary (Martin et al., 2019), which might lead manufacturers to locate themselves near places of great demand. This is especially the case when consumer preferences tend to be place specific and shaped by regional characteristics (Ormrod, 1990; Essletzbichler, 2012).

Chapter 5 contributes an important discussion to elaborate the interdependent relationship of supply and demand that has been largely neglected in EEG so far. It thereby adds to the understanding of how new industrial paths emerge in regions and emphasizes the vital role of demand in this process, which needs to be more strongly integrated in evolutionary theories of industry emergence.

6.2 Empirical contributions

6.2.1 Technology

As technologies represent the embodiment of knowledge, they were the first dimension of knowledge diffusion analyzed in this thesis. Based on their knowledge, inventors generate novel technologies that diffuse in space. How much and in which way these technologies diffuse is influenced I part by their characteristics (Pezzoni et al., 2018). One such characteristic that is relevant in this context is the level of complexity inherent in technologies (Kaufman, 1993). Although many studies emphasize the advantages of managing complex knowledge (e.g., Boltho et al., 2018; Sbardella et al., 2018), little is known about its emergence and diffusion in space. Consequently, Chapter 2 has tackled the first research question of this thesis: How do complex technologies diffuse in space?

To answer this question, Chapter 2 has combined the work of Hägerstrand (1952, 1967) with the proximity framework (Boschma, 2005) and the dimensions of technological complexity. Moreover, it has made use of an extensive data set of around four million patents that were granted in the US between 1836 and 2010. Utilizing this information, a data set has been derived for each technology considering whether a region has been granted a patent in a technology and how many years this has taken. Combined with data on geographical, technological and social relations as well as population size and patent activity, a Bayesian survival model has been calculated for each technology. Afterwards, the resulting coefficients of the independent variables have been featured in a meta-regression with technological complexity as the meta-independent variable.

The results of the investigations sustain the findings of Hägerstrand (1952): examples have been found for hierarchical diffusions in which technologies first leap to other cities and then diffuse into neighboring regions. Contagious diffusions have been observed as well, i.e., technologies diffuse from the innovator region to neighboring regions. Other technologies also leap from region to region without revealing any neighborhood effects. Crucially, the meta-regression has indicated the existence of a significant relationship between contagious diffusions and technological complexity. This adds to the work of Balland and Rigby (2017), who present evidence for such a relationship of complexity and geographic proximity. The meta-regression has also revealed that complex technologies are more quickly adopted in cities than in rural regions. This effect is even more pronounced if regions have experience in related technologies. These results complement previous findings (Balland and Rigby, 2017; Feldman et al., 2015). Interestingly, the results of social proximity vary strongly between technologies. For some, a positive relationship to adoption speed can be found. For others, this relationship is negative. Moreover, the differences in technological complexity do not explain these findings. Therefore, other reasons must be present that shall be uncovered by future research.

6.2.2 Networks

The second dimension that is of relevance when studying knowledge diffusion is networks. Social networks have been long recognized as the channels through which knowledge flows from one person to another. Hägerstrand (1965) argues that information about products is transmitted via personal communication fields. Similar, firms and regions learn about new technological opportunities through interaction in alliances or joint R&D projects (Broekel, 2015), which facilitates their innovativeness (Fornahl et al., 2011). Chapters 3 and 4 have investigated two aspects about networks that have not or only indirectly been targeted by previous research: the effectiveness of subsidized R&D networks and the simultaneous evaluation of link formation and link dissolution mechanisms.

Subsidized R&D networks

The aim of Chapter 3 was to analyze whether public induced knowledge networks lead to an increase in spatial knowledge diffusion. In contrast to most existing studies, in this work, a direct approach, similar to that of Jaffe et al. (1993), has been applied. It has evaluated whether actors’ joint participation in projects subsidized by the German government between 2000 and 2009 lead to higher patent citations among their respective regions. For the empirical analysis, a gravity model has been designed and estimated that explains inter-regional patent citations by considering relations resulting from publicly funded projects and geographical, technological and organizational proximity. In addition, relationships between regions in the form of co-inventorships have been considered.

The findings for geographical and technological proximity are in line with the existing literature as geographical proximity is negatively and technological proximity positively related to knowledge diffusion (Jaffe et al., 1993; Peri, 2005). With regards to the effects of knowledge networks, Chapter 3 presents somewhat contradictory results. The chapter has found a positive and significant relationship between co-inventors and subsequent patent citations among their respective regions. This supports the facilitating influence of networks on knowledge exchange.

In regard to policy-induced networks, the analysis has not revealed any evidence for an increased probability of patent citations between the home regions of participating organizations. This finding comes at a surprise, as, for example, Fornahl et al. (2011) and Broekel (2015) report a positive relationship between participation in subsidized joint projects and subsequent patent activity. However, in contrast to the findings of Chapter 3, these authors use an indirect approach of measuring knowledge diffusion that links project participation and patent generation but not patent citations. Accordingly, the findings in this thesis, which are based on a more direct approach, question the idea of subsidized networks facilitating knowledge flows between regions.

This may be due to a number of reasons (see Chapter 3.5). One of these may be that firms use the subsidies as windfall gains, i.e., they are not used to establish new or strengthen existing collaborations. Rather, they are simply used to reduce the firm-resources invested into existing collaborations. The results of Hagedoorn and Schackenraad (1993) support this argument. These authors conclude that organizations tend to cooperate with the same partners regardless of receiving subsidies or not. Contrastingly, Czarnitzki and Hussinger (2018) recently provided empirical evidence that private and publicly funded projects have a complementary effect on patent output. Hence, subsidized partnerships have an additional positive effect on patent outcome. Again, however, this relationship between network and patent outcome only indirectly measures whether there was a knowledge flow between partners. It is also possible that the subsidies led firms to hire additional personnel, which created new patent output. This does not mean that knowledge has been diffused between organizations. In this sense, Chapter 3 adds a methodological possibility to directly assess the effect of subsidized joint projects on knowledge flows.

The evolution of network structures

Chapter 3 has elaborated on knowledge networks and has generated new insights on their importance for inter-regional knowledge exchange. Chapter 4 has deepened the analysis and focused on the structural evolution of inter-organizational networks. More precisely, the formation and dissolution processes of a German biotech network have been analyzed. Both processes are important drivers facilitating (link formation) and hampering (link dissolution) knowledge diffusion. The determinants of link dissolution have been largely neglected in empirical papers, although they shape the structure and evolution of networks as much as formation processes (Glückler, 2007; Boschma and Frenken, 2010). Therefore, Chapter 4 has discussed and empirically explored the effects of regional characteristics, proximities and network structures on the dissolution of links in spatial knowledge networks.

Adding to the literature, the results emphasize the positive relationship of cognitive and institutional proximity with link formation (e.g., Balland, 2012; Broekel and Hartog, 2013b). Moreover, a negative relationship with geographic distance has been found, i.e., links tend to be formed between partners that are located nearby, which is also highlighted by previous studies (e.g., Ponds et al., 2007; Ter Wal, 2014). In addition, the variables at the node and structural levels are of relevance: we have observed medium and large firms to form more links than small firms, and organizations located in rural regions have been identified to be more actively forming relationships than their urban counterparts. No evidence has been found for positive effects of multi-connectivity and preferential attachment. Accordingly, organizations do not tend to link in multiple ways, and they are less likely to obtain more links if they are already well connected.

In general, the process of link dissolution seems to be more difficult to explain with factors commonly considered when studying the evolution of knowledge networks. This was shown in more variables being insignificant in the empirical models. Accordingly, fewer of them seem to be relevant for the length of subsidized joint projects in biotechnology. This observation might stem from the low variance in project length, which in turn is partly fixed by the design of the underlying policies. Nevertheless, some significant results have been discovered: the variables considering research institutes and urban locations are both negative significant implying that these organizations aim for shorter time frames. At a first glance, the latter comes as a surprise, as, according to the "third mission" of universities, they can be expected to aim for extensive knowledge exchange and project-related funding (Lee, 1996). Therefore, one could argue that they are likely to engage in longer-lasting projects. On the other hand, considering the restrictive capabilities of organizations in maintaining partnerships, the results might indicate that these institutes prefer shorter timeframes for projects to increase their chances of having many different partners and gaining access to many distinct knowledge sources (Ponds et al., 2007).

Of the considered types of proximity, only institutional proximity has become significant negative, indicating that partnerships between organizations from different institutional backgrounds, e.g., private firms and public universities, last longer. The effect contrasts with our expectations of different institutional backgrounds having no effect once the cooperation has been initialized and formally arranged. However, projects with partners of different institutional backgrounds might necessitate longer project timeframes as the involved parties need to familiarize themselves with each other and introduce joint routines and processes (Ponds et al., 2007; Balland, 2012).

6.2.3 Regional Context

Chapter 5 has studied the spatial emergence and diffusion of the German wind energy industry. More specifically, the location of new manufacturers and the location of wind turbines have been explored for the initial and growth phases of this industry. This has distinguished between local supply-push and local demand-pull mechanisms. In the first, manufacturers deploy wind turbines within their home region by creating a supportive market niche. In other words, they push their product in the market. In the latter, manufacturers emerge in regions where a significant demand already exists. So to speak, they are pulled into these regions. In order to identify which process is dominant, it has been observed and analyzed whether and when new manufacturers were founded and whether a significant relationship is given with the amount of deployed wind turbines before (demand-pull) or five years later (supply-push). Thereby, the number of future wind turbines quantifies the demand of a region.

The chapter's results, as expected, support the importance of related variety in the initial phase of the industry. Regions with former related patent portfolios are more likely to witness the emergence of wind energy manufacturers. This is in line with the findings of Neffke et al. (2011) and Breul et al. (2015). Besides this, the chapter has revealed that the emergence and diffusion of industries is also linked to demand. The likelihood to observe the foundation of wind energy manufacturers increases by 0.5% for each wind turbine that will be installed in that region in upcoming years. In the growth phase, related variety becomes less important, but demand still affects the diffusion process as distance decay effects are present. With every additional kilometer between regions and the nearest wind turbines, the likelihood of witnessing a manufacturer to emerge decreases.

Moreover, Chapter 5 confirms that the diffusion of wind turbines is strongly driven by natural conditions such as average wind speeds and available open space. Interestingly, the analysis has not found any evidence for manufacturers pushing the establishment of new wind turbines in their home region. The latter finding comes somewhat as a surprise, as it contrasts with the expectations of the transition literature (Geels and Deuten, 2006; Schot and Geels, 2008). There are a number of reasons for this. For instance, producers may not have done so because natural conditions, e.g., wind speed, were more important than proximity to their facilities. Alternatively, manufactures may have tried to push turbine deployments but were unsuccessful.

6.3 Methodological contributions

In addition to the theoretical and empirical findings, this thesis represents a substantial contribution to the methodological repertoire of evolutionary economic geography. The following section will describe these contributions and discuss their potential impact.

In Chapters 2 and 5, a Bayesian survival framework has been applied (Zhou and Hanson, 2017). Survival models focus on the explanation of event occurrences and simultaneously consider the time it has taken for the event to take place (Box-Steffensmeier and Jones, 2004). Such events can be specified as firm entries or exits, patent grants or citations. In addition, survival models allow the consideration of regional or relational characteristics influencing the likelihood of observing an event (Breul et al., 2015). The baseline hazard, i.e., the probability of an event occurring without considering any variables, can be flexibly adapted by selecting from a set of various survival models, e.g., Cox, Bayesian or Weibull (Darmofal, 2009). The spatial dependence of neighboring regions can also be taken into account and controlled for by including spatial individual or shared frailties (Ibid.). For example, Feldman et al. (2015) make use of a Cox semi-parametric survival model to investigate the spatial diffusion of biotechnology. In sum, the application of survival models appears promising for the field of economic geography, as they allow for the empirical analysis of the diffusion of knowledge in space. Thereby, they can contribute to the understanding of economic landscapes and their evolution over time, which is one of the primary aims of EEG (Boschma and Martin, 2010). In this regard, this thesis contributes to the methodological portfolio of geographers and promotes the usage of survival models for the future by offering two empirical studies as examples.

Adding to the application of a Bayesian survival framework, Chapter 2 has demonstrated the use of meta-regression analyses (Jarrell and Stanley, 1989). In order to systematically assess complex issues, economic geographers often face the challenge of how to empirically analyze the working of specific mechanisms in different contexts. In the case of innovation diffusion, such contexts can be different technologies, for which diffusion patterns and drivers might differ. Using simple regression frameworks, it is impossible to consider many of these contexts and in particular their heterogeneity. Chapter 2 has presented a solution to this issue: meta-analysis. Meta-analysis has been introduced to the scientific community to propose a quantitative method for reviewing conducted studies (Jarrell and Stanley, 1989). Typically, the variations in results of clinical studies have been evaluated with meta-analysis by considering variables that have not or could not be implemented in the single studies, e.g., group size or methods used. Relating to innovation diffusion, meta-regressions allow for factoring in an additional level (in this case, technological context). Chapter 2 has demonstrated what such a framework can look like. First, for different technologies, independent Bayesian survival models are calculated that relate regional characteristics (e.g., urbanization) and relational attributes (e.g., technological proximity) to the spatial diffusion of technologies. In a second step, the resulting coefficients are considered as the dependent variable of a supplementary regression analysis and technological characteristics, e.g., complexity, are implemented as the independent variable. Thereby, Chapter 2 contributes to the necessity of analyzing innovation diffusion in a holistic and systematical way. The possibility of recognizing the heterogeneity of contexts and analyzing this heterogeneity's impact on the underlying mechanisms presents meta-regressions as a powerful tool for economic geographers.

A further methodological contribution has been made in Chapter 4 by introducing separable temporal exponential random graph models (STERGM) to the analysis of spatial knowledge networks. STERGM is a recent advancement in the family of ERGM (Krivitsky and Handcock, 2014). ERGM allows to study the structure and evolution of networks by letting researchers consider independent variables of all three network levels: node, dyad and structural. Thus, ERGM are powerful tools already used in economic geography (e.g., Broekel and Hartog, 2013). Additionally, STERGM extends the tool by a prolific feature: the simultaneous consideration of formation and dissolution processes of links in two-mode networks. Thereby, two features of a network are addressed that have been mostly neglected in the literature thus far. First, the two-mode structure of networks is taken into consideration. Real-life networks often evolve around two-modes, e.g., projects and participants or firms and employees. Due to methodological limitations, researchers are often forced to use one-mode projections in which the first mode is dissolved, and links are created directly with the second mode (e.g., Buchmann and Pyka, 2015; Scherngell and Barber, 2009). However, this projection is likely to have questionable outcomes, as, for example, all participants of a project will have direct links in the resulting one-mode network and, thus, clustering indices tend to be very high (Liu et al., 2015). Second, STERGM differentiates between link formation and dissolution and simultaneously evaluates both processes' determinants. This is crucial as, in addition to the formation of links, their eventual dissolution is an elementary part of network evolution. However, the possibilities to analyze the latter process have been very limited so far because (a) data about link dissolution is rather rare and (b) methods that allow for disentangling the two were lacking. In Chapter 4, point (b) has been addressed by introducing STERGM to the community of economic geography. More specifically, an example has been presented that shows how to disentangle the processes of formation and dissolution and, thus, how to deepen our understanding of network evolution. Additionally, so far only a few empirical studies make use of STERGM, for example, Zhang et al. (2019).

Taken together, the individual chapters of this thesis offer new methodological approaches to the community of economic geography and thereby advance the possibilities of understanding the evolution of economic landscapes.

6.4 Limitations and future research

The eventual reflection upon this thesis shall not only bring its contributions to the scientific community to light but also its limitations. The main ones will be presented and discussed in the following section and shall serve as starting points for future research.

The evolution of complexity over time and its spatial diffusion

In Chapter 2, the diffusion of simple and complex technologies has been investigated thoroughly. The chapter thereby complements previous research (Sorenson et al., 2006; Balland and Rigby, 2017; Balland et al., 2020). However, Chapter 2 as well as the mentioned studies neglect that the complexity of a technology changes over time (Broekel, 2019). On one hand, technologies may become simpler when their components are restructured. On the other hand, economies have in general become more complex over the years (Ibid.), leading to formerly complex technologies being comparably simple today. This change in complexity is likely to affect spatial diffusion. Feldman et al. (2015) separate the diffusion of biotechnology into two phases. In the first phase, geographic proximity is unrelated to the diffusion process; in the second phase, it becomes relevant. This raises the question of whether this observation can be explained by a change in complexity. In a similar sense, Balland et al. (2013) argue that, in the video game industry, geographic proximity of interfirm collaboration becomes more important over time as complexity rises.

In accordance with these observations, a dynamic perspective on proximities might be fruitful as well to enhance our understanding of spatial diffusion processes. For example, through the constant generation and adoption of technologies, regions advance their technological portfolio. Accordingly, their technological proximity to other regions changes as well. Balland et al. (2015) identify five dynamics affecting the evolution of proximities: learning (cognitive proximity), integration (organizational proximity), decoupling (social proximity), institutionalization (institutional proximity), and agglomeration (geographic proximity). Accordingly, a further explanation for the changing role of geographic proximity in Feldman et al. (2015) may be the process of agglomeration in biotechnology (Zeller, 2001).

A dynamic perspective on complexity and proximity would allow future research to investigate how both aspects shape each other's evolution. Does the evolution of complexity shape dynamics, like organizational integration? Or do the two aspects evolve co-evolutionarily? This clearly presents a worthwhile prospect for future research.

Complexity and network structure

Chapters 3 and 4 have elaborated on the network dimension of knowledge diffusion independent of the complexity of technologies. However, both dimensions may not be independent of each other. Fleming and Sorenson (2001) study the flow of knowledge through networks and consider the level of complexity. They observe that knowledge of distinct levels of complexity diffuses differently in networks. Highly complex knowledge resists diffusion, even within closely connected clusters, whereas simple knowledge flows equally between actors, independent of their network position. This indicates that, according to its complexity, knowledge requires different network structures to flow.

Furthermore, not only may network structures shape the diffusion of complexity, but complexity may also shape the structure of networks. Economic agents that face complex technologies and, hence, greater difficulties and uncertainties may decide to interact more intensively (Carbonell and Rodriguez, 2006). In this regard, Broekel (2019) finds a positive relationship between technological complexity and number of inventors per patent. This raises the question, if not only the number of partners increases but also the time of their partnership. Does complexity affect the dissolution of network links? The interdependence between technological complexity and network evolution represents a further interesting research avenue.

The measurement of complexity

Before future research may tackle these avenues, however, a common understanding of complexity is necessary. At the moment, several definitions are at hand, leading to different measurements of complexity based on different data inputs. Fleming and Sorenson (2001) introduce "Modular Complexity" quantifying the degree of interdependence of technological subcomponents. Balland et al. (2020) measure complexity for different economic activities like scientific publications and industries. In their paper, complexity of publications is measured by the number of authors, and industries are defined as complex if their workforce requires many years of education. In another paper, Balland and Rigby (2017) compute a knowledge complexity index (KCI) by applying the complexity index developed by Hidalgo and

Hausmann (2009) to patent data. Originally this measurement was developed to evaluate countries' export and employment patterns. Just recently, Broekel (2019) developed "structural diversity" as a further index of complexity.

Comparing the empirical results of studies with different complexity measures may become difficult and may often focus exclusively on the usage of measurements. Thereby, conceptual reflection might be neglected, and theoretical progress may slow. Hence, a critical discussion on the different measures will be essential for future research. In this regard, Broekel (2019) takes a first step by comparing the results of structural diversity with modular complexity and KCI. He finds significant differences in the results, supporting the necessity of a vital discussion about the indices. Importantly, this discussion needs to tackle the question of whether there is one complexity measure that can be used in all contexts (e.g., regional and national). Or, is there a necessity for different complexity indices to analyze distinct contexts?

International knowledge diffusion

Through the course of this thesis, a regional perspective on knowledge diffusion is adopted which neglects the international level. However, today's economies are characterized by global production and knowledge networks (Binz and Truffer, 2017; Yeung and Coe, 2014). For example, the German auto manufacturer Volkswagen maintains research institutes in Germany, Spain, the US, Japan and China. Although these locations are far apart, organizational distance is likely to be relatively low, which may simplify the knowledge exchange between these facilities. In the sense of Bathelt et al. (2004), Volkswagen may provide global pipelines of knowledge, providing regions with new knowledge from outside their region.

In this regard, the literature on multi-national enterprises (MNE) appears to be a good starting point for future research (Cantwell and Iammarino, 2000; Ascani et al., 2016). For example, Schaefer and Liefner (2017) find that MNEs from developing countries can enhance their patent activities through maintaining R&D institutes abroad. This implies a knowledge flow from the R&D institutes to the headquarters. Considering these international linkages, and networks in total, would allow for the investigation of spatial diffusion patterns of technologies on a global scale. Do hierarchical and contagious patterns still dominate the diffusion of technologies, as observed in Chapter 2? Or, do new patterns occur? Are MNEs capable of transferring complex knowledge internationally?

Data, Data and again Data

The preceding paragraphs have shown that future research in economic geography faces challenging and worthwhile topics to investigate. Complexity, networks, proximities and regional contexts offer much to explore. However, all previously mentioned empirical research avenues will severely depend on the availability of data. In Chapter 4, for example, the investigation of link dissolution requires longitudinal data of inter-organizational cooperation or inter-personal partnerships that also includes the date of termination. However, this data is often not available (McPherson et al., 2001). In this thesis, we have made use of the German subsidies catalogue that includes project length, allowing us to analyze link formation as well as link dissolution. In this case, project length is defined before the project starts. Hence, the data does not capture unplanned project termination. Moreover, partners might have stopped working with each other long before the official end of the supported project.

In recent years, the enthusiasm for “big data” has grown significantly. The rapid accumulation of data through various sources like internet websites or sensors promises many new insights about all aspects of life (Labrinidis and Jagadish, 2012; Mayer-Schönberger and Cukier, 2013). Future research in economic geography might potentially benefit from this development as well by producing a more accurate and complete understanding of spatial processes (Graham and Shelton, 2013). Many data points already contain explicit or implicit spatial information (Goodchild, 2013); IP addresses include information on regional level and smartphones often track GPS positions. With respect to information on link duration and dissolution, web scraping of firm websites enables us to track the timing of formation and dissolution of partnerships. Just recently, Krüger et al. (2020) published a working paper on the “digital layer,” utilizing web-scraped information of 500,000 German firms to analyze inter-firm relations. With regards to Chapter 4, firms tend to announce strategic partnerships on their websites; if these should disappear, it might indicate the closure of the partnership. In combination with the information on projects supported by innovation policies, serving as representations of relations with initially agreed project lengths, this would allow us to investigate intended as well as unintended terminations. Additionally, in this particular use case, web scraping methods offer a time and cost effective way of acquiring data, which allows future research to analyze inter-firm networks on a large scale without limiting to specific regions or industries (Krüger et al., 2020).

6.5 Policy implications

This thesis has investigated three dimensions of knowledge diffusion: technology, networks, and regional context. Based on the empirical results, the following section will discuss the possibility of governments to intervene in markets in order to steer knowledge diffusion. In general, policy tools tackle numerous aspects of social and economic life in which market failures are perceived (Jaffe et al., 2005). Therefore, the following implications will only target a small part of potential place-based instruments.

Facilitating the diffusion of complex technologies

Chapter 2 has found evidence that complex technologies diffuse contagiously. Moreover, similar to Balland et al. (2020), our results have shown that complex technologies concentrate in urban areas. Hence, these results imply that distance dependent face-to-face interactions are essential for understanding complex technologies. Only through personal communication can complex technologies be explained and a successful transition between people be achieved (Feldman, 1993). As complex technologies promise higher economic returns, this might increase the welfare gap between regions, which would be contradictory to the aims of the EU cohesion policy (Basile et al., 2008). In order to enhance the diffusion of complex knowledge over greater distances, policy makers should facilitate the mobility of scientists and researchers. Thereby, the complex knowledge embodied in these people might flow between regions when researchers are moving between places and begin working in new regions. Either star scientists with complex knowledge enter new regions and enrich the regional stock of knowledge with more complex knowledge, or researchers acquire new knowledge by temporarily moving to regions that possess more complex knowledge. While working there they might absorb new, complex knowledge which they transmit to their home region when moving back. In order to facilitate these possibilities, policy makers could, for example, simplify the process of achieving temporary working and residence permits in their countries.

Balland et al. (2019) discuss the implications of technological complexity and relatedness on the EU smart specialization policy. They propose to identify the diversification potentials of regions by assessing the scores on relatedness and complexity to guide policy makers which technologies to support. Thereby, policy makers could facilitate the adoption of technologies that promise economic benefits and are related to regional competencies. The results of Chapter 2 also suggest that technological relatedness is beneficial when it comes to the adoption of technologies, and even more so if these technologies are complex. Therefore, policy makers should encourage firms to explore product opportunities related to their current technological portfolio but that are more complex. For example, policy makers could support firms in taking a step forward in the supply chain of their industry. Thus, instead of producing and supplying components for the next step in the production chain, they start to manufacture these as well. This would allow firms to advance, based on their skills and competencies and, therefore, increase the likelihood of a successful development.

Shaping the regional context to steer technology and industry diffusion

Current policy tools in Europe, like smart specialization, that are frequently discussed in science (Asheim et al., 2017; Balland et al., 2019; Boschma, 2014) target the supply side of regions by leveraging existing strength and facilitating the competitive advantage of regions. In Chapter 5 evidence has been revealed that demand is a further factor shaping the emergence of industries. New industries may sustain the competitiveness of regions by diversifying the regional economy (Frenken et al., 2007). Thus, policy tools may not only support supply side factors like infrastructure and human capabilities but also demand related aspects.

For governments, the most direct way to do so would be to occur as consumers themselves and create additional demand, e.g., via public tenders. As governments frequently appear to be the largest vehicle fleet and infrastructure owners, they could raise a substantial demand. In this regard, Edler and Georghiou (2007) discuss a demand-orientated innovation policy by making use of general and strategic public procurement. In case of the first option, innovation in general is an essential part of the call for tenders. The second option describes public procurement that demands specific technologies, products or services. In this way, governments could act as lead users of innovations and provide a critical mass, securing producers a certain return. Thereby, the market uncertainty accompanying the innovation process is decreased and firms might rather decide to invest in the development process. This might not only lead manufacturers to invest in production capabilities (Edler and Georghiou, 2007) but also encourage the emergence of new firms offering complementary products and services. For example, in the case of electric vehicles, the possibility that new firms target the missing charging infrastructure might be enhanced if policy makers signal a significant demand for electric vehicles.

Adjusting policy instruments according to the maturity of an industry

Similar to Neffke et al. (2011) who report that the effects of agglomeration externalities change over the industry life cycle, the results of Chapter 5 have also shown that the importance of factors vary in accordance to industry maturity. In the initial phase, related knowledge enhances the probabilities of industry emergence in a region. In the growth phase this factor loses power, and geographic proximity to demand becomes more relevant for the development of an industry. Accordingly, the results suggest that different policy tools vary in their effectiveness over the industry life cycle. Hence, policy makers should consider the maturity of an industry when designing industry supporting policies. In a similar sense, Brenner and Schlump (2011) find evidence that different policy measures affect cluster development differently in the various phases of a cluster life cycle.

In order to facilitate industry emergence and development, the results of Chapter 5 indicate that emerging industries should be supported through access to related knowledge. Policy makers could, for example, facilitate cooperation between young firms and related universities. More mature industries, on the other hand, need access to demand-related knowledge, like user experience and consumer needs, in order to perceive and understand changes in consumer preferences and react accordingly. As this kind of data likely concerns the privacy of people because it includes information about individuals' place of living or habits, the European Union is very restrictive about data capturing and storage. This requires organizations to handle a constant trade-off between data-driven innovation and privacy (Goldfarb and Tucker, 2012). Thus, privacy policy becomes successively more linked to innovation policy, which should be taken into account by policy makers when designing these instruments.

Subsidized joint projects

The subsidization of joint projects enjoys great popularity in the EU. Since 1984, the EU has established eight Framework programs with a total volume of 250 billion euros, aiming to fund joint R&D projects between organizations in different EU member states. Additionally, nations like Germany have their own funding schemes with which they support joint projects. The aim of these subsidies is to enhance the exchange of knowledge between organizations and thereby facilitate the innovation capabilities of organizations. As knowledge exchanges tend to be characterized by market failures (Lundvall, 1992), the intervention of policy makers appears legitimate (Jaffe et al., 2005). Interactive learning and innovation processes create significant costs related to invested time, effort and trust (Hagedoorn, 2002) by a simultaneous uncertainty about future returns (Bleek and Ernst, 1993). Not only may the innovation process fail, but organizations are at risk of partners behaving opportunistically (Williamson, 1973). However, while these arguments are used to justify the subsidization of joint projects, there is no empirical evidence that the amount of inter-organizational cooperation remains below a "social optimum," as one would expect because of the mentioned costs and uncertainties (Graf and

Broekel, 2020). Thus, not only does the effectiveness of this instrument need to be assessed but also its legitimation critically evaluated.

With respect to their effectiveness, Chapter 3 has analyzed whether policy-induced networks facilitate inter-regional knowledge diffusion in the form of patent citations but has failed in obtaining clear evidence as our main variable is insignificant. Several reasons appear: the knowledge transferred between partners may be obsolete and therefore does not lead to new citations because (1) novel knowledge is not existent, (2) it is (willingly) not shared, (3) it is not sharable, and (4) it cannot be shared due to partners' low absorptive capacity. With regards to (1), it might occur that novel knowledge was already shared before the induced partnership, as policy may induce similar constellations of partners in the same way as in private networks (Breschi and Cusmano, 2004; Scherngell and Barber, 2011). Therefore, it might be fruitful to try promoting unlikely partnerships, e.g., in the form of low cognitive proximity. This might also facilitate the research impact because, if successful, the partnership might create breakthrough innovations (Castaldi et al., 2015). However, the results of this thesis as well as many other studies have shown that cognitive proximity should not be too low (Nooteboom et al., 2007); otherwise a successful knowledge exchange will be highly unlikely. Thus, a balance between too low and too high cognitive proximity has to be found. In addition, policy could also aim for connecting organizations working on highly complex technologies, as these promise higher returns. However, both approaches might be likely to have a high likelihood of failure, either due to an unfavorable constellation of partner characteristics—again, low cognitive proximity tends to hamper successful cooperation—or because the confrontation with combining complex technologies might overstretch the participants' capabilities. But, if successful, the effort might be worthwhile as it could create valuable, novel technologies.

In order to enhance the possibilities of a successful knowledge exchange under these circumstances and tackle point (4), policy makers should consider utilizing different project lengths. When facing knowledge unrelated to prior experiences, organizations tend to fail to grasp the new knowledge (Cohen and Levinthal, 1990; Nooteboom et al., 2007), at least at the beginning of a project. Allowing partners to collaborate longer might neglect the problem of unrelated knowledge and enable organizations to gain new insights. Crucially, the dissolution of established links and the length of research projects are essential for successful outcomes, as different research topics and goals necessitate distinct project lengths. Accordingly, policy makers should allow individual project lengths and not restrict all projects to a certain maximum length. Today a majority of projects is limited to 36 months (see Chapter 4). But it is questionable that projects in all kinds of technologies can be completed within such a time frame. Distinguishing technologies with regards to their complexity might allow policy makers to adjust projects' lengths appropriately. Complex activities tend to take more time and resource investments (Carbonell and Rodriguez, 2006); for these, governments could define longer project lengths in contrast to simple technologies, where shorter time frames might be sufficient.

6.6 An appeal

"The history of innovations teaches us that usually it takes far too long for proven concepts and programs to become part of practice" (Glanz et al., 2008, p. 313). For example, it took the British navy until 1795 to introduce citrus juice into the diet of sailors, although its effectiveness in preventing scurvy was already recognize in 1601. Nearly two centuries were necessary to adopt a life-saving measure. Obviously, this is an example meant to exaggerate the before-mentioned quote. Nevertheless, people or regions often neglect to adopt certain processes or technologies which appear beneficial, at least from an outsider perspective. Accordingly, it is crucial to understand what factors shape the diffusion of new ideas and technologies. For economic geographers, this includes understanding the emergence and evolution of economic landscapes, and for policy makers, being able to steer socially desirable adoption processes. Therefore, at the end of this work, I would like to emphasize the importance of understanding spatial diffusion mechanisms and hope that the research in this thesis will encourage other researchers to likewise investigate spatial diffusion processes.

Bibliography

Acs, Z. J., L. Anselin, and A. Varga (2002) ‘Patents and innovation counts as measures of regional production of new knowledge’, *Research Policy*, 31, 1069–1085.

Almeida, P., and B. Kogut (1997) ‘The Exploration of Technological Diversity and Geographic Localization in Innovation: Start-Up Firms in the Semiconductor Industry’, *Small Business Economics*, 9, 21–31.

Annoni, P., and L. Dijkstra (2019) *The EU Regional Competitiveness Index 2019*., Vol. 02.

Arthur, W. (1990) ‘Positive Feedbacks in the Economy’, *Scientific American*, 262(2), 92–99.

—— (1994) *Increasing returns and path dependence in the economy*. Michigan: University of Michigan Press.

—— (2009) *The Nature of Technology: What It Is and How It Evolves*. London: Free Press.

Arundel, A., and I. Kabla (1998) ‘What percentage of innovations are patented? Empirical estimates for European firms’, *Research Policy*, 27(2), 127–141.

Ascani, A., R. Crescenzi, and S. Iammarino (2016) ‘Economic Institutions and the Location Strategies of European Multinationals in their Geographic Neighborhood’, *Economic Geography*, 92(4), 401–429.

Asheim, B., M. Grillitsch, and M. Trippl (2017) ‘Smart Specialization as an Innovation-Driven Strategy for Economic Diversification: Examples From Scandinavian Regions’. In: Radosevic S., A. Curaj, R. Gheorghiu, L. Andreescu and W. I (eds) *Advances in the Theory and Practice of Smart Specialization*, pp. 73–97. Elsevier: London.

Asheim, B. T., R. Boschma, and P. Cooke (2011) ‘Constructing Regional advantage: Platform policies based on related variety and differentiated knowledge bases’, *Regional Studies*, 45(7), 893–904.

Asheim, B. T., and M. S. Gertler (2009) ‘The Geography of Innovation: Regional Innovation Systems’, *The Oxford Handbook of Innovation*, (January 2009).

Ashwill, T. D. (2003) *Cost Study for Large Wind Turbine Blades : WindPACT Blade System Design Studies*.

Audretsch, D. B., and M. P. Feldman (1996) ‘R&D Spillovers and the Geography of Innovation and Production’, *American Economic Review*, 86(3), 630–640.

Balland, P.-A. (2012) ‘Proximity and the Evolution of Collaboration Networks: Evidence from Research and Development Projects within the Global Navigation Satellite System (GNSS) Industry’, *Regional Studies*, 46(6), 741–756.

Balland, P.-A., R. Boschma, J. Crespo, and D. L. Rigby (2019) ‘Smart specialization policy in the European Union: relatedness, knowledge complexity and regional diversification', *Regional Studies,* 53(9), 1252–1268.

Balland, P.-A., R. Boschma, and K. Frenken (2015) 'Proximity and Innovation: From Statics to Dynamics', *Regional Studies,* 49(6), 907–920.

Balland, P.-A., C. Jara-Figueroa, S. Petralia, M. Steijn, D. L. Rigby, and C. Hidalgo (2020) 'Complex Economic Activities Concentrate in Large Cities', *Nature Human Behaviour*, 4, 248–254.

Balland, P.-A., and D. Rigby (2017) 'The Geography of Complex Knowledge', *Economic Geography*, 93(1), 1–23.

Balland, P.-A., M. De Vaan, and R. Boschma (2013) 'The dynamics of interfirm networks along the industry life cycle: The case of the global video game industry, 1987-2007', *Journal of Economic Geography*, 13(5), 741–765.

Basile, R., D. Castellani, and A. Zanfei (2008) 'Location choices of multinational firms in Europe: The role of EU cohesion policy', *Journal of International Economics*, 74(2), 328–340.

Bass, F. M. (1969) 'A New Product Growth for Model Consumer Durables', *Management Science,* 15(5), 215–227.

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

Bednarz, M., and T. Broekel (2019) 'The relationship of policy induced R&D networks and inter-regional knowledge diffusion', *Journal of Evolutionary Economics*, 29, 1459–1481.

Bergek, A., and S. Jacobsson (2003) 'The emergence of a growth industry: A comparative analysis of the German, Dutch and Swedish wind turbine industries'. In: Metcalfe J Stan and U. Cantner (eds) *Change, Transformation and Development*, pp. 197–227. Physica-Verlag: Heidelberg.

Bettencourt, L. M. A., J. Lobo, D. Helbing, C. Kühnert, and G. B. West (2007) 'Growth, innovation, scaling, and the pace of life in cities', *Proceedings of the National Academy of Sciences of the United States of America*, 104(17), 7301–7306.

Binz, C., and B. Truffer (2017) 'Global Innovation Systems—A conceptual framework for innovation dynamics in transnational contexts', *Research Policy*, 46(7), 1284–1298.

Blaikie, P. (1978) 'The theory of the spatial diffusion of innovations: A spacious cul-de-sac', *Progress in Human Geography*, 2(2), 268–295.

Blaut, J. M. (1977) 'Two Views of Diffusion', *Annals of the Association of Americal*

Geographers, 67(3), 343–349.

Bleek, J., and D. Ernst (1993) *Collaborating to compete: using strategic alliances and acquisitions in the global marketplace.* New York, Chichester, Brisbane, Toronto, Singapore: John Wiley and Sons.

BMBF (2008) 'Merkblatt für Antragsteller/Zuwendungsempfänger zur Zusammenarbeit der Partner von Verbundprojekten', *Bundesministerium für Bildung und Forschung, BMBF-Vordruck 0110/10.08*.

Bode, E. (2004) 'The spatial pattern of localized R&D spillovers: An empirical investigation for Germany', *Journal of Economic Geography*, 4(1), 43–64.

Boltho, A., W. Carlin, and P. Scaramozzino (2018) 'Why East Germany did not become a new Mezzogiorno', *Journal of Comparative Economics*, 46(1), 20–34. Elsevier.

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

—— (2014) 'Constructing Regional Advantage and Smart Specialisation: Comparison of Two European Policy Concepts', *Italian Journal of Regional Science*, 13(1), 51–68.

—— (2017) 'Relatedness as driver of regional diversification: a research agenda', *Regional Studies*, 51(3), 351–364.

Boschma, R, and K. Frenken (2010) 'The spatial evolution of innovation networks: a proximity perspective'. In: Boschma Ron and R. Martin (eds) *The Handbook of Evolutionary Economic Geography*, pp. 120–135. Edward Elgar Publishing: Cheltenham.

—— (2011) 'Technological relatedness and regional branching'. In: Kogler D. F., M. P. Feldman and H. Bathelt (eds) *Beyond territory. Dynamic geographies of knowledge creation, diffusion, and innovation*, pp. 64–81. Milton Park, New York.

Boschma, R, and J. G. Lambooy (1999) 'Evolutionary economics and economic geography', *Journal of Evolutionary Economics*, 9(4), 411–429.

Boschma, R, and R. Martin (2010) 'The aims and scope of evolutionary economic geography'. In: Boschma R. and R. Martin (eds) *The Handbook of Evolutionary Economic Geography*, pp. 3–39. Edward Elgar Publishing: Cheltenham.

Boschma, R, and A. L. J. ter Wal (2007) 'Knowledge Networks and Innovative Performance in an Industrial District: The Case of a Footwear District in the South of Italy', *Industry & Innovation*, 14(2), 177–199.

Boschma, R, and R. Wenting (2007) 'The spatial evolution of the British automobile industry: Does location matter?', *Industrial and Corporate Change*, 16(2), 213–238.

Bottazzi, L., and G. Peri (2003) 'Innovation and spillovers in regions: Evidence from European patent data', *European Economic Review,* 47(4), 687–710.

Box-Steffensmeier, J., and B. Jones (2004) *Event history modeling: A guide for social scientists*. Cambridge: Cambride University Press.

Brem, A., and K. I. Voigt (2009) 'Integration of market pull and technology push in the corporate front end and innovation management-Insights from the German software industry', *Technovation*, 29(5), 351–367.

Brenner, T., and A. Mühlig (2013) 'Factors and Mechanisms Causing the Emergence of Local Industrial Clusters: A Summary of 159 Cases', *Regional Studies,* 47(4), 480–507.

Brenner, T., and C. Schlump (2011) *Policy Measures and their Effects in the Different Phases of the Cluster Life Cycle*. *Regional Studies,* Vol. 45.

Breschi, S., and L. Cusmano (2004) 'Unveiling the texture of a European Research Area : Emergence of oligarchic networks under EU Framework Programmes', *International Journal of Technology Management*, 27(8), 747–772.

Breschi, S., and C. Lenzi (2012) 'Net city: how co-invention networks shape inventive productivity in us cities', *Working Paper,* 1–32.

Breschi, S., and F. Lissoni (2005) 'Knowledge Networks from Patent Data: Methodological Issues and Research Targets'. In: Moed H., W. Glänzel and U. Schmoch (eds) *Handbook of Quantitative Science and Technology Research*, pp. 613–643. Kluwer Academic Publisher: New York, Boston, Dordrecht, London, Moscow.

—— (2009) 'Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows', *Journal of Economic Geography*, 9(4), 439–468.

Breul, M., T. Broekel, and M. Brachert (2015) 'The drivers of the spatial emergence and clustering of the photovoltaic industry in Germany', *Zeitschrift fur Wirtschaftsgeographie,* 59(3), 133–150.

Broekel, T. (2015) 'Do Cooperative Research and Development (R&D) Subsidies Stimulate Regional Innovation Efficiency? Evidence from Germany', *Regional Studies*, 49(7), 1087–1110.

—— (2019) 'Using structural diversity to measure the complexity of technologies', *Plos One,* 14(5), e0216856.

Broekel, T., and C. Alfken (2015) 'Gone with the wind? The impact of wind turbines on tourism demand', *Energy Policy,* 86, 506–519.

Broekel, T., P.-A. Balland, M. Burger, and F. van Oort (2014) 'Modeling knowledge networks in economic geography: a discussion of four methods', *The Annals of Regional Science,* 53(2), 423–452.

Broekel, T., and R. Boschma (2011) ‘Knowledge networks in the Dutch aviation industry: The proximity paradox’, *Journal of Economic Geography*, 12(2), 409–433.

Broekel, T., and T. Brenner (2011) ‘Regional factors and innovativeness: An empirical analysis of four German industries’, *Annals of Regional Science*.

Broekel, T., T. Brenner, and M. Buerger (2015) ‘An Investigation of the Relation between Cooperation Intensity and the Innovative Success of German Regions’, *Spatial Economic Analysis*, 10(1), 52–78.

Broekel, T., and H. Graf (2012) ‘Public research intensity and the structure of German R&D networks: a comparison of 10 technologies’, *Economics of Innovation and New Technology*, 21(4), 345–372.

Broekel, T., and M. Hartog (2013a) ‘Explaining the Structure of Inter- Organizational Networks using Exponential Random Graph Models’, *Industry and Innovation*, 20(3), 277–295.

—— (2013b) ‘Determinats of Cross-Regonal R&D Collaboration Networks: An Application of Exponential Random Graph Models’. In: Scherngell T. (ed.) *The Geography of Networks and R&D Collaborations*, pp. 49–70. Springer: Wien.

Brown, L., and K. R. Cox (1971) ‘Empirical Regularities in the Diffusion of Innovation’, *Annals of the Association of American Geographers*, 61(3), 551–559.

Buchmann, T., and A. Pyka (2015) ‘The evolution of innovation networks: the case of a publicly funded German automotive network’, *Economics of Innovation and New Technology*, 24(1–2), 114–139.

Buenstorf, G., M. Fritsch, and L. F. Medrano (2015) ‘Regional Knowledge, Organizational Capabilities and the Emergence of the West German Laser Systems Industry, 1975–2005’, *Regional Studies*, 49(1), 59–75.

Buisseret, T. J., H. M. Cameron, and L. Georghiou (1995) ‘What difference does it make? Additionality in the public support of R&D in large firms’, *International Journal of Technology Management*, 10(4/5/6), 587–600.

Bundesinstitut für Bau- Stadt- und Raumforschung (BBSR) (2015) ‘Laufende Raumbeobachtung – Raumabgrenzungen. Siedlungsstrukturelle Regionstypen’. Retrieved January 15, 2016, from <http://www.bbsr.bund.de/BBSR/DE/Raumbeobachtung/Raumabgrenzungen/Regionstypen/regionstypen.html?nn=443270>

Burger, M., F. van Oort, and G. J. Linders (2009) ‘On the specification of the gravity model of trade: Zeros, excess zeros and zero-inflated estimation’, *Spatial Economic Analysis*, 4(2), 167–190.

Burton, T., N. Jenkins, D. Sharpe, and E. Bossanyi (2011) *Wind Energy Handbook Second Edition*. West Sussex: John Wiley & Sons, Ltd.

Butzin, A. (2009) 'Innovationsbiographien als Methode der raum-zeitlichen Erfassung von Innovationsprozessen'. In: Dennenberg P., H. Köhler, T. Lang, J. Utz, B. Zakirova and T. Zimmermann (eds) *Innovationen im Raum - Raum für Innovationen: 11. Junges Forum der ARL, 21. bis 23. Mai 2008 in Berlin*, pp. 189–198. Verlag der ARL: Hannover.

Di Cagno, D., A. Fabrizi, V. Meliciani, and I. Wanzenböck (2016) 'The impact of relational spillovers from joint research projects on knowledge creation across European regions', *Technological Forecasting and Social Change*, 108(116), 83–94.

Camagni, R. P. (1985) 'Spatial Diffusion of Pervasive Process Innovation', *Papers of the Regional Science Association*, 58(1), 83–95.

—— (1995) 'The Concept of Innovative Milieu and Its Relevance for Public Policies in European Lagging Regions', *Papers in Regional Science*, 74(4), 317–340.

Cantner, U., and A. Meder (2008) *Regional and technological effects of cooperation behavior* (No. 14). Jenaer Economic Research Papers. Jena.

Cantwell, J., and S. Iammarino (2000) 'Multinational corporations and the location of technological innovation in the UK regions', *Regional Studies*, 34(4), 317–332.

Carbonell, P., and A. I. Rodriguez (2006) 'Designing teams for speedy product development: The moderating effect of technological complexity', *Journal of Business Research*, 59(2), 225–232.

Castaldi, C., K. Frenken, and B. Los (2015) 'Related Variety, Unrelated Variety and Technological Breakthroughs: An analysis of US State-Level Patenting', *Regional Studies*, 49(5), 767–781.

Coenen, L., P. Benneworth, and B. Truffer (2012) 'Toward a spatial perspective on sustainability transitions', *Research Policy*, 41(6), 968–979.

Cohen, L., L. Manion, and K. Morrison (2000) 'Protecting their intellectual assets: Appropriability Condtions and Why US Manufacturing Firms patent (or not)'. National Bureau of Economic Research.

Cohen, W. M., and D. A. Levinthal (1990) 'Absorptive Capacity : A New Perspective on Learning and Innovation', *Administrative Science Quarterly*, 35(1), 128–152.

Coleman, J. S. (1988) 'Social Capital in the Creation of Human Capital', *American Journal of Sociology*.

Cox, D. R. (1972) 'Regression Models and Life-Tables', *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 34(2), 187–220.

Craioveanu, M., and D. Terrell (2016) 'The impact of storms on firm survival : a bayesian spatial econometric model for firm survival'. In: Baltagi B., J. LeSage and R. Pace (eds) *Spatial Econometrics: Qualitative and Limited Dependent Variables,* Volume 37., pp. 81–118. Emerald Group Publishing Limited: Bingley.

Criscuolo, P., and B. Verspagen (2008) 'Does it matter where patent citations come from? Inventor vs. examiner citations in European patents', *Research Policy*, 37(10), 1892–1908.

Czarnitzki, D., and K. Hussinger (2018) 'Input and output additionality of R&D subsidies', *Applied Economics,* 50(12), 1324–1341.

Darmofal, D. (2009) 'Bayesian spatial survival models for political event processes', *American Journal of Political Science,* 53(1), 241–257.

DaSilva, E. (2012) 'The Colours of Biotechnology: Science, Development and Humankind', *Electronic Journal Of Biotechnology*, 7(3).

Dehmer, M., N. Barbarini, K. Varmuza, and A. Graber (2009) 'A large scale analysis of information-theoretic network complexity measures using chemical structures', *PLoS ONE*, 4(12), 20–26.

Destatis (2016) 'Daten aus dem Gemeindeverzeichnis Kreisfreie Städte und Landkreise nach Fläche und Bevölkerung auf Grundlage des ZENSUS 2011 und Bevölkerungsdichte'. Retrieved July 22, 2017, from <https://www.destatis.de/DE/ZahlenFakten/LaenderRegionen/Regionales/Gemeindeverzeichnis/Administrativ/Aktuell/04Kreise.html>

Dewald, U., and B. Truffer (2012) 'The Local Sources of Market Formation: Explaining Regional Growth Differentials in German Photovoltaic Markets', *European Planning Studies*, 20(3), 397–420.

Dodgson, M. (1992) 'The Strategic Management of R&D Collaboration', *Technology Analysis & Strategic Management*, 4(3), 227–244.

Dohse, D. (2000) 'Technology policy and the regions - The case of the BioRegio contest', *Research Policy*, 29(9), 1111–1113.

Dosi, G. (1991) 'The Research on Innovation Diffusion: An Assessment'. In: Nakićenović N. and A. Grübler (eds) *Diffusion of Technologies and Social Behavior*, pp. 179–208. Spinger: Berlin, Heidelberg.

Downs, G. W., and L. B. Mohr (1976) 'Conceptual Issues in the Study of Innovation', *Administrative Science Quarterly*, 21(4), 700–714.

Ducruet, C., and L. Beauguitte (2014) 'Spatial Science and Network Science: Review and Outcomes of a Complex Relationship', *Networks and Spatial Economics*, 14(3–4), 297–

316.

Van Eck, N. J., and L. Waltman (2009) ‘How to Normalize Co-Occurrence Data? An Analysis of Some Well-Known Similarity Measures’, *ERIM Report Series Reference No. ERS-2009-001-LIS.*

Edler, J., and L. Georghiou (2007) ‘Public procurement and innovation - Resurrecting the demand side’, *Research Policy,* 36(7), 949–963.

Emmert-Streib, F., and M. Dehmer (2012) ‘Exploring statistical and population aspects of network complexity’, *PLoS ONE,* 7(5).

Engelsman, E. C., and A. F. J. van Raan (1994) ‘A patent-based cartography of technology’, *Research Policy,* 23(1), 1–26.

Essletzbichler, J. (2012) ‘Renewable Energy Technology and Path Creation: A Multi-scalar Approach to Energy Transition in the UK’, *European Planning Studies,* 20(5), 791–816.

—— (2015) ‘Relatedness, Industrial Branching and Technological Cohesion in US Metropolitan Areas’, *Regional Studies,* 49(5), 752–766.

Etzkowitz, H., and L. Leydesdorff (2000) ‘The dynamics of innovation: From National Systems and “mode 2” to a Triple Helix of university-industry-government relations’, *Research Policy*.

Fabrizio, K., and L. Thomas (2012) ‘The impact of Local Demand on Innovation in a Global Industry’, *Strategic Management Journal*, 33, 42–64.

Feldman, M. P. (1993) ‘An Examination of the Geography of Innovation’, *Industrial and Corporate Change*, 2(1), 451–470.

Feldman, M. P., D. F. Kogler, and D. L. Rigby (2015) ‘rKnowledge : The Spatial Diffusion and Adoption of rDNA Methods’, *Regional Studies,* 49(5), 798–817. Taylor & Francis.

Fleming, L., and O. Sorenson (2001) ‘Technology as a complex adaptive system: Evidence from patent data’, *Research Policy,* 30(7), 1019–1039.

Fornahl, D., T. Broekel, and R. Boschma (2011) ‘What drives patent performance of German biotech firms? The impact of R&D subsidies, knowledge networks and their location’, *Papers in Regional Science,* 90(2), 395–418.

Fornahl, D., R. Hassink, C. Klaerding, I. Mossig, and H. Schröder (2012) ‘From the Old Path of Shipbuilding onto the New Path of Offshore Wind Energy? The Case of Northern Germany’, *European Planning Studies,* 20(5), 835–855.

Fox, J., and S. Weisberg (2011) ‘Cox Proportional-Hazards Regression for Survival Data in R’. *An Appendix to An R Companion to Applied Regression, Second Edition*. Sage Publications: Los Angeles.

Frank, O., and D. Strauss (1986) ‘Markov graphs’, *Journal of the American Statistical Association*, 81(395), 832–842.

Frenken, K., F. Van Oort, and T. Verburg (2007) ‘Related variety, unrelated variety and regional economic growth’, *Regional Studies*, 41(5), 685–697.

Friedman, T. (2005) *The World Is Flat: A Brief History of the Twenty-first Century*. New York: Farrar, Straus, and Giroux.

Garud, R., and P. Karnøe (2001) ‘Path Dependence and Creation’. In: Garud R. and P. Karnøe (eds) *Path Dependence and Creation*, pp. 1–38. Lawrence Earlbaum Associates: Mahwah.

—— (2003) ‘Bricolage vs. breakthrough: distributed and embedded ageny in technology entrepreneurship’, *Research Policy*, 32(2), 277–300.

Geels, F., and J. Deuten (2006) ‘Local and global dynamics in technological development: a socio-cognitive perspective on knowledge flows and lessons from reinforced concrete’, *Science and Public Policy*, 33(4), 265–275.

Geels, F. (2002) ‘Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study’, *Research Policy*, 31, 1257–1274.

—— (2004) ‘From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory’, *Research Policy*, 33(6–7), 897–920.

Geels, F. , and J. Schot (2007) ‘Typology of sociotechnical transition pathways’, *Research Policy*, 36(3), 399–417.

Gertler, M. (2003) ‘Tacit knowledge and the economic geography of context, or The undefinable tacitness of being (there)’, *Journal of Economic Geography*, 3, 75–99.

Giuliani, E., and M. Bell (2005) ‘The micro-determinants of meso-level learning and innovation: Evidence from a Chilean wine cluster’, *Research Policy*, 34(1), 47–68.

Glanz, K., B. K. Rimer, and K. Viswanath (2008) *Health Behaviour and Health Education*. (K. Glanz, B. K. Rimer, & K. Viswanath, Eds)*Health Education*. San Francisco: Jossey-Bass.

Glückler, J. (2007) ‘Economic geography and the evolution of networks’, *Journal of Economic Geography*, 7(5), 619–634.

—— (2010) ‘The evolution of strategic alliance network: exploring the case of stock photography’. In: Boschma Ron and R. Martin (eds) *The Handbook of Evolutionary Economic Geography*, pp. 298–315. Cheltenham.

Glückler, J., and P. Doreian (2016) ‘Editorial: social network analysis and economic geography—positional, evolutionary and multi-level approaches’, *Journal of Economic Geography*, 16, 1123–1134.

Goldfarb, A., and C. Tucker (2012) ‘Privacy and Innovation’, *Innovation Policy and the Economy*, 12(1), 65–90. University of Chicago Press.

Goodchild, M. (2013) ‘The quality of big (geo)data’, *Dialogues in Human Geography*, 3(3), 280–284.

Goodreau, S. (2007) ‘Advances in exponential random graph (p*) models applied to a large social network’, *Social Networks*, 29(2), 231–248.

Grabher, G. (1993) ‘The weakness of strong ties: the lock-in of regional development in the Ruhr area.’ In: Grabher G. (ed.) *The embedded firm. On the socioeconomics of industrial networks*. Routledge: London and New York.

Graf, H., and T. Broekel (2020) ‘A shot in the dark? Policy influence on cluster networks’, *Research Policy*, 49(3).

Graham, M., and T. Shelton (2013) ‘Geography and the future of big data, big data and the future of geography’, *Dialogues in Human Geography*, 3(3), 255–261.

Granovetter, M. (1973) ‘The Strength of Weak Ties’, *Americal Journal of Sociology*, 78(6), 1360–1380.

Granovetter, M. (1985) ‘Economic Action and Social Structure : The Problem of Embeddedness’, *The American Journal of Sociology*, 91(3), 481–510.

Griliches, Z. (1979) ‘Issues in Assessing the Contribution of Research and Development to Productivity Growth’, *The Bell Journal of Economics*.

—— (1992) ‘The Search for R&D Spillovers’. *Scandinavian Journal of Economics*, 94, 29–47.

—— (1998) ‘Patent statistics as economic indicators: a survey’. In: Elliot C. (ed.) *R&D and productivity: the econometric evidence*, pp. 287–343. University of Chicago Press: Chicago.

Grimaldi, R., and S. Torrisi (2001) ‘Codified-tacit and general-specific knowledge in the division of labour among firms: A study of the software industry’, *Research Policy*, 30(9), 1425–1442.

Grübler, A. (1997) ‘Time for a Change: On the Patterns of Diffusion of Innovation’, *Daedalus, Journal of the American Academy of Arts and Sciences*, 125(3), 19–42.

Gulati, R. (1998) ‘Alliances and Networks’, *Strategic Management Journal*, 19, 293–317.

Hagedoorn, J. (2002) ‘Inter-firm R&D partnerships: an overview of major trends and patterns since 1960’, *Research Policy*, 31(4), 477–492.

Hagedoorn, J., and J. Schackenraad (1993) ‘A comparison of private and subsidized R&D partnerships in the European Information Technology Industry’, *Journal of Common Market Studies*, 31(3), 373–390.

Hägerstrand, T. (1952) *The Propagation of Innovation Waves*. *Lund Studies in Geography, Series B Human Geography No. 4*. Lund: Lund University Press.

—— (1965) 'A Monte Carlo Approach to Diffusion', *European Journal of Sociology,* 6, 43–67.

—— (1966) 'Aspects of the Spatial Structure of Social Communication and the Diffusion of Information', *Papers of the Regional Science Association*, 16(1), 27–42.

—— (1967) *Innovation diffusion as a spatial process*. Chicago: University of Chicago Press.

Hagget, P. (2001) *Geography A Global Synthesis*. Essex: Pearson Education.

Hanneke, S., W. Fu, and E. Xing (2010) 'Discrete Temporal Models of Social Networks', *Electronic Journal of Statistics,* 4, 585–605.

Hanneke, S., and E. Xing (2007) 'Discrete temporal models of social networks'. In: Airoldi E., D. Blei, S. Fienberg, A. Goldenberg, E. Xing and A. Zheng (eds) *Statistical Network Analysis: Models, Issues, and New Directions,* pp. 115–125. Springer-Verlag: Berlin, Heidelberg.

Helfat, C., and M. Lieberman (2002) 'The birth of capabilities: Market entry and the importance of pre-history', *Industrial and Corporate Change*, 11(4), 725–760.

Herrmann, A., J. Taks, and E. Moors (2012) 'Beyond Regional Clusters: On the Importance of Geographical Proximity for R&D Collaborations in a Global Economy-the Case of the Flemish Biotech Sector', *Industry and Innovation*, 19(6), 499–516.

Hewitt-Dundas, N., and S. Roper (2010) 'Output additionality of public support for innovation: Evidence for Irish manufacturing plants', *European Planning Studies*, 18, 107–122.

Hidalgo, C. (2015) *Why Information Grows: The Evolution of Order, from Atoms to Economies*. New York: Basic Books.

Hidalgo, C., and R. Hausmann (2009) 'The building blocks of economic complexity', *Proceedings of the National Academy of Sciences*, 106(26), 10570–10575.

Hidalgo, C., B. Klinger, A.-L. Barabási, and R. Hausmann (2007) 'The product space conditions the development of nations.', *Science*, 317(5837), 482–487.

Hoekman, J., K. Frenken, and F. van Oort (2009) 'The geography of collaborative knowledge production in Europe', *Annals of Regional Science*, 43(3), 7321–7738.

Hoekman, J., T. Scherngell, K. Frenken, and R. Tijssen (2013) 'Acquisition of European research funds and its effect on international scientific collaboration', *Journal of Economic Geography*, 13(1), 23–52.

Hoover, E., and R. Vernon (1962) *Anatomy of a metropolis. The changing distribution of people and jobs within the New York metropolitan region*. Garden City, N.Y.: A Doubleday

Anchor book, A 298.

Howells, J. (2002) ‘Tacit Knowledge, Innovation and Economic Geography’, *Urban Studies,* 39(5–6), 871–884.

Hunter, D. (2007) ‘Curved Exponential Familiy Models for Social Networks’, *Social Networks,* 29(2), 216–230.

Hunter, D., S. Goodreau, and M. Handcock (2008) ‘Goodness of Fit of Social Network Models’, *Journal of the American Statistical Association*, 103(481), 248–258.

Illenberger, J., K. Nagel, and G. Flötteröd (2013) ‘The Role of Spatial Interaction in Social Networks’, *Networks and Spatial Economics,* 13(3), 255–282.

Isard, W. (1954) ‘Location Theory and Trade Theory: Short-Run Analysis’, *The Quarterly Journal of Economics*, 68(2), 305–320.

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

Jacobsson, S., and A. Johnson (2000) ‘The diffusion of renewable energy technology: An analytical framework and key issues for research’, *Energy Policy*, 28(9), 625–640.

Jacobsson, S., and V. Lauber (2006) ‘The politics and policy of energy system transformation - Explaining the German diffusion of renewable energy technology’, *Energy Policy,* 34(3), 256–276.

Jaffe, A. (1986) ‘Technological Opportunity and Spillovers of R & D: Evidence from Firms’ Patents, Profits, and Market Value’, *The American Economic Review,* 76(5), 984–999.

Jaffe, A., R. Newell, and R. Stavins (2005) ‘A tale of two market failures: Technology and environmental policy’, *Ecological Economics*, 54, 164–174.

Jaffe, A., and M. Trajtenberg (1999) ‘International Knowledge Flows: Evidence From Patent Citations’, *Economics of Innovation and New Technologies*, 8, 105–136.

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

Jarrell, S., and T. Stanley (1989) ‘Meta-regression analysis: A quantitative method of literature surveys’, *Journal of Economic Surveys,* 3(2), 161–170.

Jensen, M., B. Johnson, E. Lorenz, and B. Lundvall (2007) ‘Forms of knowledge and modes of innovation’, *Research Policy*, 36(5), 680–693.

Johnson, A., and S. Jacobsson (2003) ‘The emergence of growth industry: a comparative analysis of the German, Dutch and Swedish wind turbine industries’. In: Metcalfe J.S. and U. Cantner (eds) *Change, Transformation and Development,* pp. 197–227. Physica-Verlag: New York.

Justman, M. (1994) 'The Effect of Local Demand on Industry Location', *The Review of Economics and Statistics*, 76(4), 742–753.

Kaldellis, J., and D. Zafirakis (2011) 'The wind energy (r)evolution: A short review of a long history', *Renewable Energy*, 36(7), 1887–1901.

Kaldor, N. (1970) 'The Case fo Regional Policies*', *Scottish Journal of Political Economy*, 17(3), 337–348.

Kammer, J. (2011) *Die Windenergieindustrie. Evolution von akteuren und Unternehmensstrukturen in einer Wachstumsindustrie mit rämlicher Perspektive*. (F. N. Nagel, Ed.). Hamburg: Geographische Gesellschaft Hamburg, Franz Steiner Verlag.

Kaufman, S. (1993) *The Origins of Order: Self organisation and selection in evolution*. New York, Oxford: Oxford University Press.

Kirkland, E. (1961) *Industry Comes of Age: Business, Labor, and Public Policy, 1860-1897*. Chicago: Quadrangle Books.

Klepper, S. (1996) 'Entry, Exit, Growth, and Innovation Over the Product Life Cycle', *The American Economic Review*, 86(3), 562–583.

—— (1997) 'Industry Life Cycles', *Industrial and Corporate Change*, 6(1), 145–182.

—— (2006) 'The Evolution of Geographic Structure in New Industries', *Revue de l'OFCE*, 5, 136–158.

—— (2007) 'Disagreements, Spinoffs, and the Capital of the U.S. Automobile Industry', *Management Science*, 53(4), 616–631.

Kline, S., and N. Rosenberg (1986) 'An Overview of Innovation', *European Journal of Innovation Management*, 38, 275–305.

Klitkou, A., and L. Coenen (2013) 'The Emergence of the Norwegian Solar Photovoltaic Industry in a Regional Perspective', *European Planning Studies*, 21(11), 1796–1819.

Knoben, J., and L. Oerlemans (2006) 'Proximity and inter-organizational collaboration: A literature review', *International Journal of Management Reviews*, 8(2), 71–89.

Kogut, B., W. Shan, and G. Walker (1992) 'The make-or-cooperate decision in the context of an industry network'. In: Nohiria N. and R. Eccles (eds) *Networks and organizations*, pp. 348–365. Harvard Business School Press: Brighton.

Kosfeld, R., and A. Werner (2012) 'Deutsche Arbeitsmarktregionen – Neuabgrenzung nach den Kreisgebietsreformen 2007–2011', *Raumforschung und Raumordnung*, 70, 49–64.

Krivitsky, P., and S. Goodreau (2015) 'STERGM - Separable Temporal ERGMs for modeling discrete relational dynamics with statnet'.

Krivitsky, P., and M. Handcock (2014) 'A separable model for dynamic networks', *Journal of* *the Royal Statistical Society. Series B: Statistical Methodology*, 76(1), 29–46.

Krüger, M., J. Kinne, D. Lenz, and B. Resch (2020) *The Digital Layer: How Innovative Firms Relate on the Web*. *ZEW Discussion Paper*.

Krugman, P. (1991) 'Increasing returns and economic geography', *Journal of Political Economy*, 99(3), 483–499.

Kulke, E. (2006) *Wirtschaftsgeographie*., 2nd ed. Paderborn: Ferdinand Schöningh.

Labrinidis, A., and H. Jagadish (2012) 'Challenges and opportunities with big data', *Proceedings of the VLDB Endowment*, 5(12).

Lee, Y. (1996) '"Technology transfer" and the research university: A search for the boundaries of university-industry collaboration', *Research Policy*, 25(6), 843–886.

Leifeld, P., and S. Cranmer (2015) 'A theoretical and empirical comparison of the temporal exponential random graph model and the stochastic actor-oriented model', *Network Science*, 7(1), 20–51.

Lewis, J., and R. Wiser (2007) 'Fostering a renewable energy technology industry: An international comparison of wind industry policy support mechanisms', *Energy Policy*, 35(3), 1844–1857.

Liu, X., B. Derudder, and Y. Liu (2015) 'Regional geographies of intercity corporate networks: The use of exponential random graph models to assess regional network-formation', *Papers in Regional Science*, 94(1), 109–126.

Liu, X., B. Derudder, Y. Liu, F. Witlox, and W. Shen (2013) 'A stochastic actor-based modelling of the evolution of an intercity corporate network', *Environment and Planning A*, 45(4), 947–966.

Loasby, B. (2001) 'Organisation as interpretative systems', *Revue d'économie industrielle*, 97, 17–34.

Lufin Varas, M. (2007) *Essays in social space: Applications to Chilean communities on inter -sector social linkages, social capital, and social justice*. *ProQuest Dissertations and Theses*.

Lundvall, B. (1992) *National systems of innovation: Towards a theory of innovation and interactive learning*. London: Pinter Publishers.

Lundvall, B., and B. Johnson (1994) 'The learning economy', *Journal of Industry Studies*, 1(2), 23–42.

Maggioni, M., M. Nosvelli, and T. Uberti (2007) 'Space versus networks in the geography of innovation: A European analysis', *Papers in Regional Science*, 86(3), 471–493.

Maggioni, M., T. Uberti, and M. Nosvelli (2014) 'Does intentional mean hierarchical?

Knowledge flows and innovative performance of European regions', *Annals of Regional Science*, 53, 453–485.

Maggioni, M., T. Uberti, and S. Usai (2011) 'Treating patents as relational data: Knowledge transfers and spillovers across Italian provinces', *Industry and Innovation*, 18(1), 39–67.

Makino, S., C. Chan, T. Isobe, and P. Baemish (2007) 'Intended and Unintended Termination of International Joint Ventures', *Strategic Management Journal*, 28, 1113–1132.

Marshall, A. (1920) *Principles of economics. An introductory volume.*, 8th ed. London: Macmillan.

Martin, H., R. Martin, and E. Zukauskaite (2019) 'The multiple roles of demand in new regional industrial path development: A conceptual analysis', *Environment and Planning A: Economy and Space*, 51(8), 1741–1757.

Mayer-Schönberger, V., and K. Cukier (2013) *Big Data. A Revolution that will Transform how we Live, Work and Think*. London: John Muray.

McPherson, M., L. Smith-lovin, and J. Cook (2001) 'Birds of a Feather: Homophily in Social Networks', *Annual Review of Sociology*, 27, 415–444.

Menzel, M., and D. Fornahl (2009) 'Cluster life cycles-dimensions and rationales of cluster evolution', *Industrial and Corporate Change*, 19(1), 205–238.

Meyer-Krahmer, F. (1985) 'Innovation Behaviour and Regional Indigenous Potential', *Regional Studies*, 19(6), 523–534.

Mills, M. (2011) 'The fundamentals of survival and event history analysis'. In: Mills M. (ed.) *Introducing Survival and Event History Analysis*, pp. 1–17. Sage Publications: London.

Montanaria, A., and A. Saberi (2010) 'The spread of innovations in social networks', *Proceedings of the National Academy of Sciences of the United States of America*, 107(47), 20196–20201.

Montresor, S., and F. Quatraro (2019) 'Green technologies and Smart Specialisation Strategies: a European patent-based analysis of the intertwining of technological relatedness and key enabling technologies', *Regional Studies*.

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

Morris, M., M. Handcock, and D. Hunter (2008) 'Specification of Exponential-Family Random Graph Models: Terms and Computational Aspects.', *Journal of statistical software*, 24(4), 1548.

Morrison, A., S. Petralia, and D. Diodato (2018) *Migration and invention in the age of mass migration. Papers in Evolutionary Economic Geography (PEEG)*.

Murphy, J. (2003) 'Social space and industrial development in East Africa: deconstructing the logics of industry networks in Mwanza, Tanzania', *Journal of Economic Geography*, 3(2), 173–198.

Myrdal, G. (1957) *Economic Theory and Underdeveloped Regions*. London: Harper Torchbooks.

Neffke, F., M. Henning, and R. Boschma (2011) 'How Do Regions Diversify over Time? Industry Relatedness and the Development of New Growth Paths in Regions', *Economic Geography*, 87(3), 237–265.

Neffke, F., M. Henning, R. Boschma, K. Lundquist, and L. Olander (2011) 'The dynamics of agglomeration externalities along the life cycle of industries', *Regional Studies*, 45(1), 49–65.

Nelson, R., and S. Winter (1982) *An Evolutionary Theory of Economic Change*. Cambridge: Belknap Press.

Neukirch, M. (2010) *Die internationale Pionierphase der Windenergienutzung*. Georg-August-Universität Göttingen.

Nooteboom, B., W. Van Haverbeke, G. Duysters, V. Gilsing, and A. van den Oord (2007) 'Optimal cognitive distance and absorptive capacity', *Research Policy*, 36(7), 1016–1034.

North, D. C. (1955) 'Location Theory and Regional Economic Growth', *Journal of Political Economy*, 63(3), 243–258.

Ohlhorst, D. (2009) *Windenergie in Deutschland. Konstellationen, Dynamiken und Regulierungspotenziale im Innovationsprozess. Mit einem Geleitwort von Prof. Dr. Martin Jänicke.* Freie Universität Berlin.

Opsahl, T. (2013) 'Triadic closure in two-mode networks: Redefining the global and local clustering coefficients', *Social Networks*, 35(2), 159–167.

Ormrod, R. (1990) 'Local Context and Innovation Diffusion in a Well-Connected World', *Economic Geography*, 66(2), 109–122.

Paci, R., and S. Usai (2009) 'Knowledge flows across European regions', *Annals of Regional Science*, 43(3), 669–690.

Paier, M., and T. Scherngell (2011) 'Determinants of collaboration in European R&D networks: Empirical evidence from a discrete choice model', *Industry and Innovation*, 18(1), 89–104.

Park, J., and M. Newman (2004) 'Statistical mechanics of networks', *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics,* 70(6).

Parkhe, A. (1991) 'Interfirm Diversity, Organizational Learning, and Longevity in Global

Strategic Alliances', *Journal of International Business Studies*, 22, 579–601.

Pavitt, K. (1998) 'Technologies, products & organization in the innovating firm: What Adam Smith tells us and Joseph Schumpeter doesn't.', *Industrial and Corporate Change*, 7(3), 433–452.

Peri, G. (2005) 'Determinants of Knowledge Flows and Theirs Effect on Innovation', *Review of Economics and Statistics*, 87(2), 308–322.

Perkins, R., and E. Neumayer (2005) 'The international diffusion of new technologies: A multitechnology analysis of latecomer advantage and global economic integration', *Annals of the Association of American Geographers*, 95(4), 789–808.

Petralia, S., P.-A. Balland, and D. Rigby (2016) 'Unveiling the geography of historical patents in the United States from 1836 to 1975', *Scientific Data*, 3, 1–14.

Pezzoni, M., R. Veugelers, and F. Visentin (2018) 'Is This Novel Technology Going to Hit?', *Academy of Management Proceedings*, 2018(1), 13832.

—— (2019) 'How fast is a novel technology going to be a hit? Antecedents predicting follow-on inventions', *Academy of Management Proceedings*, 2019(1), 13223.

Phene, A., K. Fladmoe-lindquist, and L. Marsh (2006) 'Breakthrough Innovations in the U.S. Biotechnology Indusrty: The Effects of Technological Space and Geographic Origin', *Strategic Management Journal*, 27, 369–388.

Polanyi, M. (1966) *The Tacit Dimension*. New York: Doubleday.

Polidoro, F. J., G. Ahuja, and W. Mitchell (2011) 'When the Social Structure Overshadows Competitive Incentives: The Effects of Network Embeddedness on Joint Venture Dissolution', *Academy of Management Journal*, 54(1), 203–223.

Ponds, R., F. van Oort, and K. Frenken (2010) 'Innovation, spillovers and university-industry collaboration: An extended knowledge production function approach', *Journal of Economic Geography*, 10(2), 231–255.

Ponds, R., F. Van Oort, and K. Frenken (2007) 'The Geographical and Institutional Proximity of Research Collaboration', *Papers in Regional Science*, 86(3), 423–443.

Porter, M. (1990) 'The Competitive Advantage of Nations', *Havard Business Review*, 1(1), 73–91.

—— (1998) *On competition*. Boston: The Harvard business review book series.

Powell, W., K. Koput, and L. Smith-Doerr (1996) 'Interorganizational Collaboration and the Locus of Innovation: Networks of learning in biotechnology', *Administrative Science Quarterly*, 41(1), 116–145.

Pugliese, E., G. Chiarotti, A. Zaccaria, and L. Pietronero (2017) 'Complex economies have a lateral escape from the poverty trap', *PLoS ONE*, 12(1), 1–18.

Pyke, F., G. Becattini, and W. Sengenberger (1990) *Industrial districs and inter-form co-operation in Itlay*. Geneva: International Institute for Labour Studies.

Robins, G., P. Pattison, Y. Kalish, and D. Lusher (2007) 'An introduction to exponential random graph (p*) models for social networks', *Social Networks*, 29(2), 173–191.

Rodríguez-Pose, A., and R. Crescenzi (2008) 'Mountains in a flat world: Why proximity still matters for the location of economic activity', *Cambridge Journal of Regions, Economy and Society*, 1(3), 371–388.

Rogers, E. (2003) *Diffusion of innovations*., 5th ed. New York: Free Press.

Romer, P. (1990) 'Endogenous Technological Change', *Journal of Political Economy*, 98(5), 72–102.

Saxenian, A. (1994) *Regional advantage. Culture and competition in Silicon Valley and Route 128*., 2nd ed. Cambridge: Harvard University Press.

Sbardella, A., E. Pugliese, A. Zaccaria, and P. Scaramozzino (2018) 'The role of complex analysis in modeling economic growth', *Entropy*, 20(11).

Schaefer, K., and I. Liefner (2017) 'Offshore versus domestic: Can EM MNCs reach higher R&D quality abroad?', *Scientometrics*, 113(3), 1349–1370.

Scherngell, T., and M. Barber (2009) 'Spatial interaction modelling of cross-region R&D collaborations: Empirical evidence from the 5th EU framework programme', *Papers in Regional Science*, 88(3), 531–546.

—— (2011) 'Distinct spatial characteristics of industrial and public research collaborations: Evidence from the 5th EU Framework Programme', *Annals of Regional Science*, 46(2), 247–266.

Schilling, M., and C. Phelps (2007) 'Interfirm collaboration networks: The impact of large-scale network structure on firm innovation', *Management Science*, 53(7), 1113–1126.

Schlaile, M., K. Bogner, and M. Müller (2018) 'Knowledge diffusion in formal networks: the roles of degree distribution and cognitive distance', *International Journal of Computational Economics and Econometrics*, 8(3/4), 388.

Schmoch, U., and F. Laville (2003) 'Linking technology areas to industrial sectors. Final Report to the European Commission'.

Schot, J., and F. Geels (2008) 'Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy', *Technology Analysis & Strategic Management*, 20(5), 537–554.

Schrader, S. (1991) 'Informal technology transfer between firms: Cooperation through information trading', *Research Policy*, 20(2), 153–170.

Schwartz, M., F. Peglow, M. Fritsch, and J. Günnther (2012) 'What drives innovation output from subsidized R & D cooperation ?— Project- level evidence from Germany', *Technovation*, 32(6), 358–369.

Scott, A., and M. Storper (1986) *High technology industry and regional development: a theoretical critique and reconstruction*. Berkshire: Geographical Papers - University of Reading, Department of Geography.

Shah, S., and M. Tripsas (2007) 'The accidental entrepreneur: the emergent and collective process of user entrepreneurship', *Strategic Entrepreneurship Journal*, 1(1–2), 123–140.

Short, L. (2002) 'Wind Power and English Landscape Identity'. In: Pasqualetti M., P. Gipe and R. Righter (eds) *Wind Power in View: Energy Landscapes in a Crowded World*. Academic Press: San Diego.

Simmie, J., R. Sternberg, and J. Carpenter (2014) 'New technological path creation: evidence from the British and German wind energy industries', *Journal of Evolutionary Economics*, 24(4), 875–904.

Simon, H. (1962) 'The Architecture of Complexity', *Proceedings of the American Philosophical Society*, 106(10), 467–482.

Singh, K. (1997) 'The Impact of Technological Complexity and Interfirm Cooperation on Business Survival', *The Academy of management Journal*, 40(2), 339–367.

Sorenson, O., J. Rivkin, and L. Fleming (2006) 'Complexity, networks and knowledge flow', *Research Policy*, 35(7), 994–1017.

Sternberg, R. (2000) 'Innovation networks and regional development-evidence from the European Regional Innovation Survey (ERIS): Theoretical concepts, methodological approach, empirical basis and introduction to the theme issue', *European Planning Studies*, 8(4), 389–407.

—— (2003) 'Wissensintensität und regionales Umfeld als Determinanten der Entstehung und Entwicklung junger Unternehmen'. In: Steinle C. and K. Schumann (eds) *Gründung von Technologieunternehmen*, pp. 219–237. Gabler Verlag: Wiesbaden.

Stevens, F. (1926) *The Beginnings of the New York Central Railroad: A History*. New York: G. P. Putnam's Sons.

Storper, M., and A. Venables (2004) 'Buzz: Face-to-face contact and the urban economy', *Journal of Economic Geography*, 4(4), 351–370.

Storper, M., and R. Walker (1989) *The capitalist imperative; territory, technology and industrial growth*. New York: Basil Blackwell.

De Tarde, G. (1903) *The laws of imitation*. New York: H. Holt and Company.

Teece, D. (1981) 'The Market for Know-How and the Efficient International Transfer of Technology', *The Annals of the American Academy of Political and Social Science*, 458(1), 81–96.

Tether, B. (2002) 'Who co-operates for innovation, and why - an empirical analysis', *Research Policy*, 31(6), 947–967.

Therneau, T., C. Crowson, and E. Atkinson (2017) 'Using Time Dependent Covariates and Time Dependent Coefficients in the Cox Model Time', *Red*, 1–16.

Theyel, G. (2012) 'Spatial Processes of Industry Emergence : US Wind Turbine Manufacturing', *European Journal of Futures Research*, 50(5), 857–870.

Thompson, W. (1972) 'The National System of Cities as an Object of Public Policy', *Urban Studies*, 9(1), 99–116.

Tijssen, R. (1998) 'Quantitative assessment of large heterogeneous R&D networks: The case of process engineering in the Netherlands', *Research Policy*, 26(7–8), 791–809.

Tinbergen, J. (1962) *Shaping the World Economy*. New York: Twentieth Century Fund.

Truffer, B., and L. Coenen (2012) 'Environmental Innovation and Sustainability Transitions in Regional Studies', *Regional Studies*, 46(1), 1–21.

Tsouri, M. (2019) 'Knowledge transfer in time of crisis: evidence from the Trentino region', *Industry and Innovation*, 26(7), 820–842.

Usselman, S. (1991) 'Patents Purloined : Railroads , Inventors , and the Diffusion of Innovation in 19th-Century America', *Technology and Culture*, 32(4), 1047–1075.

Verspagen, B. (1997) 'Measuring Intersectoral Technology Spillovers: Estimates from the European and US Patent Office Databases', *Economic Systems Research*, 9(1), 47–65.

Vinciguerra, S., K. Frenken, and M. Valente (2010) 'The geography of internet infrastructure: An evolutionary simulation approach based on preferential attachment', *Urban Studies*, 47(9), 1969–1984.

Ter Wal, A. (2014) 'The dynamics of the inventor network in German biotechnology: geographic proximity versus triadic closure', *Journal of Economic Geography*, 14(3), 589–620.

Ter Wal, A., and R. Boschma (2009) 'Applying social network analysis in economic geography: framing some key analytic issues', *The Annals of Regional Science*, 43(3), 739–756.

Wanzenböck, I., T. Scherngell, and R. Lata (2015) 'Embeddedness of European Regions in European Union-Funded Research and Development (R&D) Networks: A Spatial

Econometric Perspective', *Regional Studies*, 49(10), 1685–1705.

Watts, D., and S. Strogatz (1998) 'Collective dynamics of "small-world" networks.', *Nature*, 393(6684), 440–442.

Weber, A. (1909) *Über den Standort der Industrien*. Tübingen: Verlag von J. C. B. Mohr.

Wiederaufbau, K. für (2012) 'KMU-Definition. Allgemeine Erläuterungen zur Definition der Kleinstunternehmen so-wie der kleinen und mittleren Unternehmen (KMU)'. Retrieved January 30, 2016, from <https://www.kfw.de/Download-Center/Förderprogramme-(Inlandsförderung)/PDF-Dokumente/6000000196-KMU-Definition.pdf>

Williamson, O. (1973) 'Markets and hierarchies, some elementary considerations', *American Economic Review*, 63(2), 316–325.

Witt, U., T. Broekel, and T. Brenner (2012) *Knowledge and its economic characteristics: A conceptual clarification* (No. 2007–013). Jena Economic Research Paper.

Yeung, H. W. Chung, and N. Coe (2014) 'Toward a Dynamic Theory of Global Production Networks', *Economic Geography*, 91(1), 29–58.

Young, H. P. (2000) *The Diffusion of Innovations in Social Networks* (No. 437). Economics Working Paper Archive.

Zander, U., and B. Kogut (1995) 'Knowledge and the Speed of the Transfer and Imitation of Organizational Capabilities: An Empirical Test', *Organisation Science*, 6(1), 76–92.

Zeller, C. (2001) 'Clustering Biotech: A Recipe for Success? Spatial Patterns of Growth of Biotechnology in Munich, Rhineland and Hamburg', *Small Business Economics*, 17, 123–141.

Zhou, H., T. Hanson, and J. Zhang (2016) 'spBayesSurv: Fitting Bayesian Spatial Survival Models Using R', *Journal of Statistical Software*, VV(II), 1–33.

Zhou, H., and T. Hanson (2017) 'A Unified Framework for Fitting Bayesian Semiparametric Models to Arbitrarily Censored Survival Data, Including Spatially Referenced Data', *Journal of the American Statistical Association*, 113(522), 571–581.

Zucker, L., M. Darby, and J. Armstrong (1998) 'Geographically localized knowledge: Spillovers or markets?', *Economic Inquiry*, 36(1), 65–86.

Nederlandse samenvatting

De EU Regional Competitiveness Index 2019 laat opnieuw zien dat het economische concurrentievermogen in Europa per regio significant verschilt (Annoni and Dijkstra, 2019). Voor deze index wordt gekeken naar regionale kenmerken zoals de kwaliteit van de infrastructuur, het opleidingsniveau van de beroepsbevolking en het aantal octrooiaanvragen. Op die schaal blijken tussen regio's grote verschillen te bestaan, waarbij Stockholm het hoogst scoort en regio's zoals het noordelijke Egeïsche gebied in Griekenland het slechtst. Eén aspect van het regionale concurrentievermogen dat in de literatuur wordt benadrukt is het vermogen om vaak te innoveren, aangezien innovaties regio's in staat stellen om nieuwe inkomsten en welvaart te genereren (Porter, 1990). Maar gezien de cumulatieve aard van kennisgeneratie is het vermogen om innovatie te creëren veelal ruimtelijk geconcentreerd in een aantal regio's (Feldman, 1993; Acs et al., 2002; Balland, 2017). Dit leidt ertoe dat de economische verschillen tussen regio's groter worden (Rodríguez-Pose and Crescenzi, 2008).

Een vitale en wederzijdse kennisoverdracht tussen regio's zou de verkleining van deze economische verschillen kunnen ondersteunen. De uitwisseling van kennis zou regio's in staat kunnen stellen om zelf innovaties te produceren (Asheim and Gertler, 2009) of in ieder geval economisch gebruik te maken van innovaties, al hebben ze deze niet zelf geproduceerd. Maar: "de geschiedenis van de innovatie leert ons dat het meestal veel te lang duurt voordat bewezen concepten en programma's onderdeel worden van de praktijk" (Glanz et al., 2008: 313). Evenals de productie van innovaties lijkt de verspreiding hiervan ruimtelijk gebonden te zijn en de neiging te hebben om verspreiding te weerstaan (Audretsch and Feldman, 1996; Jaffe et al., 1993). Daarom wordt in dit proefschrift onderzocht waarom sommige regio's veel sneller kennis overnemen dan andere en worden de factoren die een rol spelen bij de ruimtelijke verspreiding van kennis geanalyseerd.

De verspreiding van innovaties is een onderzoeksonderwerp in verschillende vakgebieden, waaronder economie, sociologie, marketing en innovatieonderzoek (e.g. De Tarde, 1903; Bass, 1969; Griliches, 1992; Rogers, 2003). Het is een belangwekkend onderwerp vanwege de gevolgen van deze verspreiding voor het economisch handelen en het maatschappelijk leven. Gewoontes en routines veranderen niet door het genereren van innovatie, maar door de daaropvolgende brede verspreiding ervan. Zo zat er bijvoorbeeld 190 jaar tussen de ontdekking dat met citrusvruchtensap scheurbuik te voorkomen was en de formele opname hiervan in het voedingspatroon van zeelieden. Toen pas veranderden de omstandigheden aan boord van

Engelse schepen (Glanz et al., 2008). Maar ongeacht het belang hiervan hebben economisch geografen hun aandacht de afgelopen jaren niet op de verspreiding van innovaties gericht, maar op de productie ervan. Dat was een aanvullende reden om in dit proefschrift de mechanismen van innovatieverspreiding te analyseren en ons begrip daarvan te vergroten.

In dit proefschrift worden drie hoofddimensies van ruimtelijke verspreiding geïdentificeerd die van invloed zijn op de wijze waarop innovaties worden verspreid. De eerste hier geanalyseerde dimensie is technologie, oftewel de concretisering van kennis. Technologieën zijn het resultaat van het onderzoeken, combineren en testen van nieuwe kenniscombinaties (Arthur, 2009). De kenmerken van deze gecombineerde kenniscomponenten hebben vervolgens gevolgen voor de verspreiding van technologie (Rogers, 2003). Zo kan de kennis die in technologie besloten ligt zowel gecodeerd als onuitgesproken zijn (Polanyi, 1966; Nelson and Winter, 1982; Gertler, 2003). Dat laatste belemmert de verspreiding significant, doordat de kennis daardoor niet kan worden uitgeschreven en gemakkelijk overgedragen. Meestal is er veelvuldige interactie tussen economische actoren op persoonlijk niveau voor nodig om dergelijke kennis uit te wisselen. Die interacties maken deel uit van de tweede dimensie van kennisverspreiding waarnaar in dit proefschrift wordt gekeken, namelijk netwerken. Hägerstrand (1966) benadrukte dat de verspreiding van innovaties klaarblijkelijk de sociale relaties tussen mensen volgt. En ook bij economisch-geografisch onderzoek wordt er sterk gekeken naar de mechanismen van netwerkontwikkeling en kennisuitwisseling (Glückler, 2007; Ter Wal and Boschma, 2009; Boschma and Frenken, 2010). Daarbij ligt de nadruk sterk op het relationele niveau en de effecten van nabijheid op de vorming van netwerkverbindingen (Boschma, 2005; Knoben and Oerlemans, 2006; Ter Wal, 2014). Naast de twee genoemde dimensies wordt in dit proefschrift nog een derde onderzocht: de regionale context. De regionale context geeft veelal vorm aan de perceptie die mensen hebben met betrekking tot welke soorten innovaties waardevol zijn om over te nemen (Blaut, 1977; Ormrod, 1990). Innovaties worden in een specifieke context ontwikkeld en de succesvolle verspreiding ervan is grotendeels afhankelijk van de vraag of potentiële gebruikers deze kunnen benutten in hun lokale context. Op de volgende pagina’s zullen deze dimensies, de bijbehorende lacunes in onderzoek en de relevante empirische hoofdstukken van dit proefschrift in meer detail worden beschreven.

Technologie

Over het algemeen kunnen technologieën worden opgevat als systemen van onderling met elkaar verbonden subcomponenten (Arthur, 2009). Dergelijke componenten zijn direct én indirect met elkaar verbonden en vormen als zodanig een ‘combinatorisch netwerk’ (Broekel, 2019). Die netwerken kunnen eenvoudig of complex van structuur zijn. Complexe technologieën worden gekenmerkt door een relatief groot aantal componenten met een groot aantal onderlinge verbindingen. Daarom hebben mensen veel informatie nodig om complexe technologieën te kunnen doorgeven en begrijpen (Simon, 1962). Daar komt nog bij dat de onderliggende kennis vaker onuitgesproken zal zijn (Broekel, 2019). Daardoor is het begrijpen en stimuleren van complexe technologieën lastiger en bewerkelijker. Maar juist doordat deze moeilijk te imiteren en te kopiëren zijn, bieden ze vaak aanzienlijke concurrentievoordelen en economische waarde (Balland and Rigby, 2017; Sbardella et al., 2018).

Hoewel de voordelen en knelpunten van complexiteit in de literatuur worden benoemd, bestaat er maar weinig empirisch onderzoek waarin de verspreiding van complexe technologieën wordt geanalyseerd. Een uitzondering wordt gevormd door Sorenson et al. (2006), Feldman et al. (2015) as well as Balland and Rigby (2017), die de verspreiding van complexe technologieën bekijken aan de hand van de dimensies van nabijheid (Boschma, 2005). En niet alleen is het empirisch onderzoek dun gezaaid, de bestaande onderzoeken laten ook nog eens verschillende resultaten zien, zoals in dit proefschrift wordt aangetoond. Daarom wordt in hoofdstuk 2 de ruimtelijke verspreiding geanalyseerd van enkele honderden technologieën waarop in de 19e en 20e eeuw in de VS octrooi is verleend. Om preciezer te zijn: het werk van (Hägerstrand, 1952; 1966; 1967) op het gebied van ruimtelijke verspreidingspatronen, de dimensie van technologische complexiteit (Simon, 1962; Kaufman, 1993) en het concept nabijheid (Boschma, 2005) worden hier gecombineerd. Alle drie deze concepten kunnen nuttig zijn voor onderzoekers om inzicht te krijgen in ruimtelijke verspreidingsprocessen, maar zijn vooralsnog nooit in onderling verband geanalyseerd. Daarom wordt in hoofdstuk 2 onderzocht of de ruimtelijke verspreidingspatronen van eenvoudige en complexe technologieën van elkaar verschillen. Naast de hiërarchische en ‘besmettelijke’ verspreidingspatronen die door Hägerstrand (1967) zijn geïntroduceerd, wordt in hoofdstuk 2 ook gekeken naar verspreidingsvormen waarbij technologieën van regio naar regio ‘springen’, naast de effecten van technologische en sociale nabijheid.

Om de gevolgen van deze factoren op de ruimtelijke verspreiding te analyseren en te beoordelen of deze per complexiteitsniveau verschillen is gebruikgemaakt van een procedure die uit twee stappen bestaat. Voor de eerste stap zijn 285 Bayesiaanse overlevingsmodellen doorgerekend (Zhou and Hanson, 2017), één voor elke technologie, met inachtneming van geografische, technologische en sociale nabijheid evenals de regionale diversiteit wat betreft populatie en technologie. Voor de tweede stap zijn de afgeleide coëfficiënten gebruikt om verschillende metaregressies te modelleren (Jarrell and Stanley, 1989) met technologische complexiteit als meta-afhankelijke variabele. Dat heeft het mogelijk gemaakt om in hoofdstuk 2 empirisch te onderzoeken of de effectgrootte van de bekeken variabelen verandert met het niveau van complexiteit. Verschilt bijvoorbeeld het effect van technologische nabijheid op de verspreiding tussen eenvoudige en complexe technologieën?

Uit de resultaten komt naar voren dat technologieën significant van elkaar verschillen wat betreft ruimtelijke verspreiding. In dat verband zijn voorbeelden van hiërarchische, besmettelijke en springende verspreiding gevonden. Bovendien wordt de kans om een besmettelijk verspreidingspatroon waar te nemen groter naarmate de complexiteit toeneemt. Complexe technologieën hebben aldus de neiging om zich in golfachtige patronen vanuit hun oorsprongsregio te verspreiden naar aangrenzende regio's, en van daaruit naar verdere aangrenzende regio's. In hoofdstuk 2 wordt ook duidelijk dat technologische nabijheid tussen de regio van de innovator en de regio van de gebruiker nog gunstiger is voor complexe technologieën.

Netwerken

In hoofdstuk 3 en 4 worden de twee lacunes onderzocht die zijn geïdentificeerd met betrekking tot de netwerkdimensie van kennisverspreiding. In dat verband wordt in hoofdstuk 3 aandacht besteed aan de tweede lacune en de volgende onderzoeksvraag: faciliteren gesubsidieerde gezamenlijke R&D-projecten de verspreiding van kennis tussen regio's die aan dezelfde gezamenlijke R&D-projecten deelnemen?? Daarna volgt hoofdstuk 4, waarin wordt onderzocht of verschillende vormen van nabijheid van invloed zijn op de desintegratie van verbindingen binnen kennisnetwerken.

De afgelopen jaren is er veel onderzoek verricht waarbij de effecten van netwerken op de overloop van kennis empirisch werden geanalyseerd. Daarbij brengen veel onderzoeken een positief effect aan het licht van interorganisatorische relaties op de innovatieactiviteiten van organisaties en regio's (Hewitt-Dundas en Roper 2010; Maggioni et al. 2014; Broekel 2015). In lijn met deze bevindingen zijn ook beleidsmakers de nadruk gaan leggen op interorganisatorische relaties. Zowel de Europese Unie als natiestaten, waaronder Duitsland, hebben programma's geïnitieerd waarmee zij gezamenlijke onderzoeksprojecten subsidiëren. De effecten van deze programma's zijn eveneens door onderzoekers beoordeeld. Een recente voorbeeld is het werk van Schwartz et al. (2012), Di Cagno et al. (2016) as well as Czarnitzki and Hussinger (2018), die positieve effecten op de innovatieprestaties van organisaties aantonen. Bij al deze onderzoeken wordt echter een empirische analyse gemaakt van het verband tussen deelname aan gezamenlijke projecten en de daaropvolgende innovatieoutput. Men gaat er dus van uit dat er kennisuitwisseling heeft plaatsgevonden als de innovatieoutput na afloop van het project toeneemt. Op die manier is het positieve effect van gesubsidieerde R&D-projecten op kennisverspreiding een kwestie van interpretatie en blijft onduidelijk of er kennis is uitgewisseld tussen partners.

In hoofdstuk 3 wordt deze lacune in onderzoek geanalyseerd op basis van de benadering van Jaffe et al. (1993), die octrooivermeldingen als ‘papieren spoor’ van kennisverspreiding hebben geïntroduceerd. Aan de hand daarvan worden de effecten van gesubsidieerde gezamenlijke projecten op interregionale octrooivermeldingen onderzocht. Om precies te zijn wordt er een zwaartekrachtmodel voor de gegevens ontwikkeld, waarbij het aantal vermeldingen tussen twee regio’s als afhankelijke variabele wordt genomen en projectdeelname (drie respectievelijk vijf jaar eerder) als onafhankelijke variabelen. Daarnaast wordt in hoofdstuk 3 gekeken naar verdere relationele en regionale kenmerken, zoals technologische nabijheid en de regionale octrooiproductie. Interessant genoeg wordt in hoofdstuk 3 geen significant bewijs gevonden voor een effect van gesubsidieerde R&D-projecten op daaropvolgende octrooivermeldingen. Dat kan verschillende oorzaken hebben, die in hoofdstuk 3 grondig worden behandeld.

De derde lacune heeft betrekking op de vraag of nabijheden een verklaring kunnen zijn voor de desintegratie van netwerkverbanden. De laatste jaren is er veel aandacht geweest voor de vorming van netwerkverbanden en de bijbehorende mechanismen als verklaring voor de evolutie van interorganisatorische netwerken (e.g. Murphy, 2003; Boschma and Ter Wal, 2007; Broekel and Boschma, 2011; Ter Wal, 2014). Uit dit onderzoek komt naar voren dat de verschillende dimensies van nabijheid significante effecten hebben op de vorming van netwerkverbanden. Zo heeft Scherngell and Barber (2009) bijvoorbeeld het effect geanalyseerd van geografische en technologische nabijheid op de vorming van verbanden en daarbij een positieve relatie ontdekt tussen beide nabijheidsdimensies en samenwerking op het gebied van R&D.

Maar ondanks de zeer omvangrijke literatuur over netwerkevolutie is de tegenhanger van netwerkevolutie – de desintegratie van verbanden – grotendeels genegeerd. En dat terwijl netwerkevolutie en “[…] -variatie moeten worden beschouwd als resultaten van endogene mechanismen van netwerkvorming en -desintegratie” (Glückler, 2007: p. 627). Daarnaast vormt de desintegratie van verbanden een grote belemmering voor kennisverspreiding, aangezien organisaties die hun banden verbreken stoppen met het uitwisselen van kennis. Daarom wordt in hoofdstuk 4 deze lacune in de literatuur onderzocht op basis van een nieuwe methodologische benadering. Onlangs hebben Krivitsky and Handcock (2014) zogenaamde scheidbare temporele exponentiële random graafmodellen (STERGM) ontwikkeld, waarmee de vorming en desintegratie van netwerkverbanden gelijktijdig kunnen worden geanalyseerd. In hoofdstuk 4 worden gegevens over gesubsidieerde gezamenlijke projecten in de Duitse biotechnologische industrie geschikt gemaakt voor dit kader door variabelen af te leiden op de niveaus van knooppunt, dyade en netwerkstructuur.

Uit de analyse komt naar voren dat de effecten van de beschouwde variabelen met betrekking tot verbandvorming en -desintegratie varieert, wat de stelling ondersteunt dat deze processen significant van elkaar verschillen. Hoofdstuk 4 vormt een aanvulling op de bestaande empirische literatuur over verbandvorming doordat er significante verbanden worden gevonden tussen geografische, cognitieve en institutionele nabijheid enerzijds en verbandvorming anderzijds. Daarnaast blijkt dat institutionele nabijheid een positieve relatie vertoont met de desintegratie van verbanden. Dat duidt erop dat organisaties met een soortgelijke organisatorische achtergrond – bijvoorbeeld twee non-profitorganisatie – verbanden sneller verbreken dan organisaties met verschillende achtergronden.

Regionale context

De derde dimensie van kennisverspreiding waaraan aandacht wordt besteed is de regionale context. De afgelopen decennia is er regelmatig gediscussieerd over de vraag of moderne informatie- en communicatietechnologieën, zoals het internet, de wereld zouden 'afvlakken' met betrekking tot innovatievermogen en concurrentievoordeel (Friedman, 2005). Het betoog van Feldman (1993) gaat nog steeds op: regio's hebben verschillende capaciteiten en daardoor uiteenlopende mogelijkheden om nieuwe kennis en technologieën te produceren (Rodríguez-Pose and Crescenzi, 2008; Balland et al., 2020). Naast kennisproductie is ook het gebruik van technologieën een plaatsgebonden proces, betogen Blaut (1977) en Ormrod (1990). De beoordeling of een technologie of product de moeite van het onderzoeken en testen waard is, is afhankelijk van de regionale omstandigheden waarin mensen leven. Als zij besluiten om gebruik te maken van nieuwe technologieën, zoals windenergie, kunnen er nieuwe bedrijfstakken ontstaan. De eerste gebruikers kunnen een bepaalde technologie verbeteren, deze aan andere aanbieden en op die manier handel creëren (Shah and Tripsas, 2007). In dergelijke gevallen ontstaat er een bedrijfstak naar aanleiding van vraag. Vraag is dan ook een belangrijke drijfveer voor het ontstaan van nieuwe bedrijfstakken (Martin et al., 2019). Binnen de economische geografie is echter veel meer gekeken naar de aanbodzijde om het ontstaan en de ontwikkeling van bedrijfstakken te verklaren – denk aan padafhankelijkheid en gerelateerde variëteit – dan naar de vraag (Garud and Karnøe, 2001; Frenken et al., 2007; Boschma and

Frenken, 2011). Daarom wordt in hoofdstuk 5 van dit proefschrift de volgende onderzoeksvraag behandeld: geeft de lokale vraag vorm aan de ruimtelijke opkomst van bedrijfstakken?

Gegevens over de Duitse windenergie-industrie van 1983 tot 2010 zijn geschikt gemaakt voor twee Bayesiaanse overlevingskaders. Aan de ene kant wordt in hoofdstuk 5 geanalyseerd of, en hoe snel, regio's windenergie-installaties in gebruik hebben genomen door te onderzoeken of er sprake is van een stimulans door het aanbod ('supply push'). Aan de andere kant wordt er onderzocht of, en hoe snel, er nieuwe aanbieders van windenergie zijn opgericht. Daarbij wordt gekeken of er sprake is van aanzuiging vanuit de vraag ('demand pull'), waarbij de geplande ingebruikname van windenergie-installaties als verklarende factor wordt gehanteerd. Naast deze variabelen wordt in de modellen ook rekening gehouden met regionale factoren en omgevingsfactoren, zoals de aanwezigheid van gerelateerde bedrijfstakken en de gemiddelde windsnelheid.

De resultaten vormen een aanvulling op het bestaande onderzoek door de nadruk te leggen op het belang van gerelateerde variëteit, verstedelijking en industriële agglomeratie voor het ontstaan van bedrijfstakken. Daarnaast levert hoofdstuk 5 ook belangrijk empirisch bewijs voor het belang van vraag voor de locatie van nieuwe bedrijven.

Conclusie

De in dit proefschrift gepresenteerde bevindingen vormen een essentiële volgende stap in het economisch-geografisch onderzoek, doordat er vier belangrijke lacunes in onderzoek worden onderzocht. Daarnaast is het onderwerp van het onderzoek – kennisverspreiding – een onderwerp waaraan in de academische gemeenschap onvoldoende aandacht wordt besteed. Daarom wordt in dit proefschrift een innovatieve behandeling van de theoretische benadering van dit onderwerp gepresenteerd, waarbij de literatuur over complexiteit, nabijheid en ruimtelijke verspreidingspatronen met elkaar worden gecombineerd. Bovendien worden er innovatieve methodes en empirische settings gebruikt om de gestelde onderzoeksvragen te beantwoorden. Bayesiaanse overlevingsmodellen, STERGM en meta-analyse worden geïntroduceerd en met succes toegepast op de gegevens. In het laatste hoofdstuk worden beleidsimplicaties uitgewerkt op basis van de empirische bevindingen. Daarnaast worden in dit hoofdstuk nieuwe onderzoeksrichtingen voor de toekomst gepresenteerd die worden geïmpliceerd door de bevindingen en beperkingen van het hier gepresenteerde werk. Zo vormt bijvoorbeeld de onderlinge afhankelijkheid van technologische complexiteit en netwerkevolutie een interessant onderzoeksgebied voor de toekomst. Hier moet met klem worden opgemerkt dat de onderzochte dimensies van kennisverspreiding elkaar op geen enkele wijze uitsluiten en er waarschijnlijk nog andere mechanismen voor ruimtelijke verspreiding zijn die kunnen worden onderzocht.

Acknowledgments

When I first came into touch with the idea of doing a PhD, I was fully concerned with writing my master thesis and parallelly being a working student at Mercedes-Benz Vans. So, I was not really thinking of writing a PhD thesis. In that moment, my former team leader approached me with the question of whether I would not be interested in writing my PhD and thereby continuing to work for the company. Therefore, my first thanks go to Ralf Kehrberger and Renate Reichenauer who initiated this process and gave me the chance to obtain my doctoral degree by offering me the opportunity to stay at that company while also being able to keep on researching.

With this chance and idea in mind, I have approached Tom Broekel who already supervised my master thesis at that point and asked whether he could imagine supervising my PhD in the industry. Typically Tom, he was directly on board but nevertheless mentioned that it depended on me getting my master's degree first. Like I didn't know that. Anyway, my utter most gratitude goes to him as a person, supervisor and academic role model. Tom is always full of ideas which clearly supported me in finding my own way. More than that he has always been understanding and supportive regarding my full employment position at Mercedes-Benz Vans that has forced me to write my PhD at the weekends. But nevertheless, he has always been demanding the highest scientific quality in designing and conducting research and, thereby, I have learned a lot. Without his passionate support, I am confident that I would not have been able to finish this thesis. Thank you for your support, your advice and your continuous pushing me further. After we both had moved to Utrecht University, Ron Boschma and Andrea Ascani began supervising me as well. With them, they brought fresh ideas and new perspectives to my work that enriched it significantly. Therefore, I want to give many thanks to you.

My PhD was not only influenced by my academic fellows, but also by family and friends. They have always supported me when I was stuck by pulling me away from the desk to get a free mind. In this regard, I am most thankful for my fiancée Victoria. I know that she would have loved us two to explore South Germany, meet friends more often or just go for a leisurely brunch in one of Stuttgart's many cafes. However, as the mornings are my best and productive time of the weekends, this rarely happened. Nevertheless, she has always fully supported me in writing and finishing my thesis. She has been there when I wanted to throw everything away, cheered me up, and gave me new strength to proceed. More than that, she has also been an inspiration and a wonderful sparring partner challenging my ideas, thoughts and arguments. I will be always thankful for her and I am looking forward to a future with her and with many more shared experiences to come. Starting with the movement to our own home and our wedding to happen this summer. I love you!

I also want to thank my parents. They have not only given me the opportunity to learn and be educated but they always encouraged me to strive for more and pushed me a little bit further. They have also respected my decision to study Geography, although they have clearly asked themselves, what job I will find with it. Nevertheless, they respected my decision and believed in me. Thank you!

Many more people have been there for me over the last years, and although I have not mentioned you all by name, be sure that I have not forgotten you and I am very thankful for all the people I have met, the conversations I had and the experiences we shared. And I am looking forward for all the new ones to come.

Curriculum vitae

Marcel Bednarz was born on February 18, 1990 in Bielefeld, Germany. He lives and works in Stuttgart, Germany. He holds a bachelor's degree of Geography from the Westfälische Wilhelms-Universität Münster (2013) and a Master of Economic Geography from the Leibniz Universität Hannover (2016). While studying in Hannover he did an exchange semester at the Universiteit Utrecht (2014). In 2016, his master thesis was awarded as the best of the year. Afterwards, Marcel Bednarz started to work at Daimler AG in Stuttgart as a PhD student before becoming a full-time employee in 2017.