institute for economic research and policy

#2205 Bremen Papers on Economics & Innovation

Transforming Regional Knowledge Bases: A Network and Machine Learning Approach to Link Entrepreneurial Experimentation and Regional Absorptive Capacity

Jessica Birkholz*

April 2022

Abstract

This study explores the regional innovation system characteristics that build the basis for the regional absorptive capacity of entrepreneurial knowledge. Regionalized patent data is combined with firm level and regional information for German regions over the period 1995 until 2015. Network analysis is applied to identify regional innovation system characteristics on three different layers: 1) cooperation between incumbent firms, 2) learning regimes, and 3) the technological knowledge base. Random forest analyses on basis of conditional inference classification trees are used to identify the most important characteristics for the regional absorption of entrepreneurial knowledge in general and on different efficiency levels. It is shown that characteristics on all three layers impact the regional absorption of entrepreneurial knowledge. Further, the direction and magnitude of the effect regional innovation system characteristics have on the regional knowledge absorption vary across different levels of absorption rates. It is concluded that for a successful implementation of policies to increase the impact of entrepreneurial knowledge on regional development, the regional innovation system needs to be monitored and adapted continuously.

Keywords

Entrepreneurship; Regional absorptive capacity; Smart specialization

JEL Classifications

L26; O33; D85

*University of Bremen, Faculty of Business Studies and Economics
Max-von-Laue-Straße 1,
28359 Bremen, Germany
jessica.birkholz@uni-bremen.de
ORCID: 000-0003-4252-2184

Universität Bremen

1. Introduction

Knowledge is a critical success factor for regional economic performance (Caragliu & Nijkamp, 2012). In dynamic economic environments and in times of ever growing globalization, fast technological progress, and changing demand patterns, existing knowledge can become obsolete quickly (Neffke et al., 2018). This results in deprivation of regional competitive advantages (Neffke et al., 2018; Tushman & Anderson, 1986). To counteract this, a modern innovation policy has to fulfil two objectives: 1) fostering the generation of knowledge and 2) fostering the absorption and implementation of the new knowledge into the knowledge base, which turns it into a potential competitive advantage (Qian & Jung, 2017; Uyarra, 2019). Thus, the generation of new knowledge is necessary but not sufficient to build regional competitive advantages. System-wide innovative activities are required to transform knowledge into an advantage (Caragliu & Nijkamp, 2012).

This needs to be considered by regional innovation policies, such as the smart specialisation strategy (S3). A key element in this increasingly important policy approach is the 'entrepreneurial discovery', which is seen as the missing link between knowledge and economic growth (Braunerhjelm et al., 2010; D'Adda et al., 2019; Foray et al., 2012). According to this perspective, entrepreneurs are, due to their involvement into the development of technologies, in the position to not only create and use knowledge but also to introduce dynamism into the regional economy (Carlsson et al., 2013; D'Adda et al., 2019; Malerba & McKelvey, 2020; Schumpeter, 1934).

Former research finds that the effects of entrepreneurship on the regional economy, i.e. the symbiosis of entrepreneurial and incumbent firms, are conditioned by characteristics of regional innovation systems. Thus, the same entrepreneurship may be in the position to kick off system-wide innovative activities in one region but not in another (Fernandez-Serrano et al., 2019; Szerb et al., 2019). This underlines the importance of a regionally embedded view on entrepreneurship and its effects, because a universal, well-defined relationship between firm foundation rates, innovative activities, and economic growth is not found in empirical studies (Fritsch, 2008). Therefore, regional absorption mechanisms – the identification, assimilation and exploitation of knowledge – are a key factor to be considered when designing policies which aim at the development of regional knowledge bases via entrepreneurial discovery processes.

The overall rational of policies such as the S3, i.e. the exploitation of regional strengths for competitive advantages, can only be fulfilled with an in-depth understanding of mechanisms in regional innovation systems (Uyarra, 2019). However, it has been criticized that these policies and scientific research do not pay enough attention to the diffusion of the knowledge generated by entrepreneurial experimentation (Lindholm-Dahlstrand et al., 2019; Uyarra, 2019). Despite the broad application of the absorptive

capacity concept and acknowledgement of its multi-dimensional character, so far researchers mostly looked at it as a static resource rather than a multi-layered capability (Camisón & Forés, 2010; Choi & Park, 2017; Distel, 2019; Molina-Morales et al., 2019; Russo-Spena & Di Paola, 2019; Schmidt, 2005). Empirical studies frequently measure absorptive capacity through R&D activities (Cohen & Levinthal, 1990). This only takes into account a fraction of the absorptive capacity construct. For the purpose of designing effective policy interventions, a deeper understanding of regional innovation system characteristics that build regional absorptive capacities is inevitable.

These considerations call for an integrated perspective on entrepreneurship and regional absorptive capacity. This study aims to contribute to the understanding on how to improve regional absorptive capacities for entrepreneurial knowledge (Lane et al., 2002; Schmidt, 2005). Moreover, this study addresses the calls from former researchers to analyze the micro-level processes that shape structural change (Boschma, 2017; Grillitsch, 2019; Uyarra, 2010) and its effects on the regional level (Pihlajamaa, 2018; Schmidt, 2005; Szerb et al., 2019).

The purpose of this study is to analyze the effect and the importance of the regional innovation system characteristics for the regional absorptive capacity of entrepreneurial knowledge. The present study analyses German regions in the time frame from 1995 until 2015. The applied data set consists of regionalized patent data combined with firm level and regional information. Specifics of regional innovation systems are assessed by network indicators on three different layers that enable identification, assimilation, and exploitation of knowledge: 1) cooperation between incumbent firms, 2) learning regimes, and 3) the technological knowledge base. These are subsequently implemented into two random forest analyses based on conditional inference classification trees and correlational analyses to answer three overarching research questions:

RQ1: Which regional innovation system characteristics enabling identification, assimilation, and exploitation of knowledge are the most important for absorbing entrepreneurial knowledge?

RQ2: Which regional innovation system characteristics enabling higher efficiency in identification, assimilation, and exploitation of knowledge are the most important for absorbing entrepreneurial knowledge?

RQ3: Which effect does these important regional innovation system characteristics have on different levels of entrepreneurial knowledge absorption?

The main contribution of this study is twofold. First, it is shown that regional absorptive capacity is a multi-dimensional concept. Key factors for the absorption of entrepreneurial knowledge are identified in each layer of the regional innovation system - cooperation between incumbent firms, learning regimes, and technological knowledge bases. The

results suggest that regional innovation system characteristics have a heterogeneous association to regional absorption rates across different levels of absorption of entrepreneurial knowledge. Second, this study shows the explanatory power of machine learning techniques in innovation system research by allowing complex interactions and high correlation between independent variables. The approach taken by this study enables to analyze regional innovation systems in a holistic way. Identifying the driving factors of networks, which build absorptive capacities, would not be possible with commonly used multivariate statistical methods. While multivariate methods reduce the complexity of innovation systems to a set of simple relationships, the approach taken in this study considers the complex nature of regional innovation systems.

The remainder of this paper is organized as follows. The second chapter discusses the interplay of entrepreneurial experimentation, regional innovation systems, and regional absorptive capacity. In Chapter three the patent, firm, and regional level data base of this study is described. Chapter four presents the methodological approach with focus on random forest analyses on basis of conditional inference classification trees. The results are outlined in Chapter five and discussed in Chapter six.

2. State of the art

2.1. Regional innovation policies and entrepreneurial experimentation

Scientific debates emphasize the importance of knowledge and innovation as key drivers of structural change, economic development, and competitive advantages (Ferreira Moutinho, 2016; Grillitsch, 2019; Lindholm-Dahlstrand et al., 2019; Lopez-Bazo & Motellon, 2018; Teirlinck & Spithoven, 2019). It is argued that competitive advantages do not depend any longer on natural resources or other traditional inputs, but on creation of new knowledge and ideas as well as their incorporation (David & Foray, 2002; Ferreira Moutinho, 2016; Furman & Hayes, 2004). There are two main reasons to view knowledge as a regional resource for competitive advantages. First, innovative economic activities are unequally distributed, and geographical clustering of knowledge-intensive activities results in self-sufficient regional innovation systems (Asheim & Gertler, 2006; Teirlinck & Spithoven, 2019). Second, tacit knowledge is most commonly shared via personal contacts, and these are assumed to be facilitated by spatial proximity, which enhances the capacity of localized learning (Asheim & Gertler, 2006; Asheim & Isaksen, 2002; Bathelt et al., 2004; Storper & Venables, 2004; Teirlinck & Spithoven, 2019). Therefore, it is proposed that the subnational level is appropriate to capture the knowledge dynamics (Acs & Armington, 2004). The regional knowledge base, which accumulates itself in the long term, lies the foundation for subsequent innovative activities (Balland & Rigby, 2017). For this reason, the development of knowledge bases is targeted by innovation policy (D'Adda et al., 2019).

One of the most notable policy-related conceptual advancements is the S3 that was a radical novelty in EU regional policy (Pyka et al., 2019). The smart specialisation approach is assumed to create more diversity among regions by focussing on already existing technological domains in which superior innovative capabilities can be found and diversification of these into related sectors (D'Adda et al., 2019; Foray, 2018). By acknowledging the systemic nature of innovation, the S3 aims at the transformation of the economic structures of a region (Foray, 2018; Pyka et al., 2019). One main novelty is that it follows a bottom-up approach, i.e. an 'entrepreneurial discovery process' guides the transformation of the regional knowledge base (D'Adda et al., 2019; Foray et al., 2011; Uyarra, 2019). Therefore, entrepreneurs are of key importance for knowledge creation and diffusion and, thus, regional development (Braunerhjelm et al., 2010). It is argued that entrepreneurs have superior knowledge to choose the most promising domains for future development (D'Adda et al., 2019). Here, the focus lies on innovative entrepreneurs, directly involved in the development of technologies. (D'Adda et al., 2019). A regional innovation system relies on entrepreneurs who look for a way to exploit knowledge, which is left unused by incumbents (Acs et al., 2009; Neffke et al., 2018; Qian & Jung, 2017). This is further supported by the finding that entrepreneurs are especially important to generate growth from unrelated variety in a region (Fritsch & Kublina, 2018). With increasingly more complex innovation systems and exponentially rising opportunities for knowledge recombination, this entrepreneurial function becomes a main driver of economic development (Grillitsch, 2019; Schumpeter, 1934). Even more importantly, entrepreneurial experimentation is required to kick off system-wide activities from incumbent firms (Lindholm-Dahlstrand et al., 2019). Thus, the transformative capabilities stemming from entrepreneurs are an essential element for structural change in regions (Grillitsch, 2019).

Although this process of entrepreneurial discovery and the transformation of regional knowledge bases can occur spontaneously and in a decentralized manner, policy can aim at maximising its impact (Foray, 2018). By diverging from conventional regional policies that were funding industries and technologies, which did not require a detailed investigation of the regional diversification potential, policy-makers are confronted with new challenges. Policy-makers who want to implement the rational of the smart specialisation need to consider the regional knowledge base to design new entrepreneurship policy interventions instead of replicating successful policies issued in other regions (Foray, 2018; Mason & Brown, 2014; Spigel, 2017; Szerb et al., 2019).

2.2. Constitution of the regional absorptive capacity

To be able to design an efficient regional innovation policy, which builds upon entrepreneurial experimentation as well as innovative activities performed by incumbent firms, it is of crucial importance to understand the dynamics between entrepreneurial and incumbent firms.

Empirical research struggles to find a well-specified relationship between the emergence of new firms, innovative activities, and economic growth (Fritsch, 2008). A reason might be that the transformative capabilities of entrepreneurial experimentation only come into play when the regional innovation system absorbs the newly generated knowledge and, consequently, kicks off system-wide experimentations including incumbent firms (Caragliu & Nijkamp, 2012; Foray, 2018; Lindholm-Dahlstrand et al., 2019).

The link between entrepreneurial experimentation and regional knowledge base transformation is the regional absorptive capacity (Qian & Jung, 2017). Absorptive capacity is the ability to *"identify, assimilate and exploit knowledge from the environment"* (Cohen & Levinthal, 1989, p. 569). Despite several modifications of this concept[1], common ground is that absorptive capacity is a multi-dimensional construct (Schmidt, 2005). The concept of absorptive capacity is not limited to its original firm level perspective but can be transferred to the regional level (Schmidt, 2005). For the prosperity of regional innovation systems and to create competitive advantages over other regions, it is of key importance to identify, assimilate, and exploit newly generated knowledge stemming from intraregional entrepreneurial experimentation. Unexploited knowledge spills over to other regions when the regional innovation system is not suited to absorb it appropriately (Caragliu & Nijkamp, 2012).

Entrepreneurial and incumbent firms are required to have a symbiotic relationship in order to form a well-functioning innovation system (Lindholm-Dahlstrand et al., 2019). Although entrepreneurial firms can initiate short-run structural change in an innovation system, the persistence of this change is limited due to high failure rates of these firms (Neffke et al., 2018). Further, less access to capital markets, difficulties in acquiring external financial resources, and the lack of human capital prevent entrepreneurial firms from exploiting the full commercial potential of their generated knowledge (Carpenter & Petersen, 2002; Colombo & Grilli, 2007; Lindholm-Dahlstrand et al., 2019). In contrast to this, incumbent firms can rely on their accumulated knowledge and experiences in the market and, thus, can scale up the application of knowledge (Lindholm-Dahlstrand et al., 2019). However, these firms avoid exploring new knowledge but rather refine their accumulated base (Arvanitis & Woerter, 2015; Neffke et al., 2018; Pihlajamaa, 2018; Rothaermel & Deeds, 2004). In conclusion, entrepreneurial and incumbent firms have respective advantages at different stages of the innovation process and, therefore, take on complementary roles in an innovation system (Baumol, 2002; Lindholm-Dahlstrand et al., 2019). Due to the division of tasks within the innovation process, in which entrepreneurial ventures produce radical new knowledge and incumbent firms absorb this and develop it further, knowledge production (i.e. the generation of new knowledge) and knowledge absorption (i.e. the identification, assimilation, and exploitation of existing knowledge) are two distinct processes. These processes are conducted by different actors and follow different regularities. Thus, knowledge absorption can be understood

[1]For an extensive overview see Zahra & George, 2002.

as a coordination mechanism between different actors in an innovation system (Pihlajamaa, 2018). Following these considerations, the first proposition is stated as follows:

Proposition 1: Entrepreneurial knowledge production does not equal entrepreneurial knowledge absorption in regard to underlying mechanisms.

2.3. Regional absorptive capacity and regional innovation system characteristics

The regional absorptive capacity is likely to be influenced by context specifics (Pihlajamaa, 2018). Since the absorptive capacity is built in a path-dependent manner, contextual differences can exist in how information and knowledge is transferred in a region (Cantner & Graf, 2008; Cohen & Levinthal, 1990). This context specificity can be consulted to explain different impacts of the same type of entrepreneurship in two distinct innovation systems (Fernandez-Serrano et al., 2019; Teirlinck & Spithoven, 2019). Thus, the effect of entrepreneurial experimentation depends on the framework conditions of the regional innovation system, which are governing the process of regional knowledge absorption.

To analyze these characteristics, regional innovation systems can be broken down into three parts – a set of nodes, the relations between these nodes, and formal or informal rules which determine the interactions (Pyka et al., 2019). Through this system of nodes and relations, knowledge is transferred or constructed among heterogeneous actors (Boschma et al., 2014). This means, regional absorptive capacity requires collective action. In this regard the collective action governs the efficiency of the absorption process. These considerations lead to the second and third proposition:

Proposition 2: Regional innovation system characteristics are important for the absorption of entrepreneurial knowledge in general.

Proposition 3: Regional innovation system characteristics are important for the efficiency of the absorption of entrepreneurial knowledge.

However, because of the complex nature of regional innovation systems, it is unlikely that the indicators have a linear relation to the innovation outcome of the system (Pyka et al., 2019). Absorption rates differ between regions, within these different levels of absorption rates it is likely that the characteristics of the regions differ. Therefore, it is hypothesized that regions with the same level of knowledge absorption have the same characteristics in common. Thus, the fourth proposition is stated as follows:

Proposition 4: The relation of regional innovation system characteristics to the regional absorption of entrepreneurial knowledge differs between levels of absorption rates.

3. Data

3.1. Creation of the data set

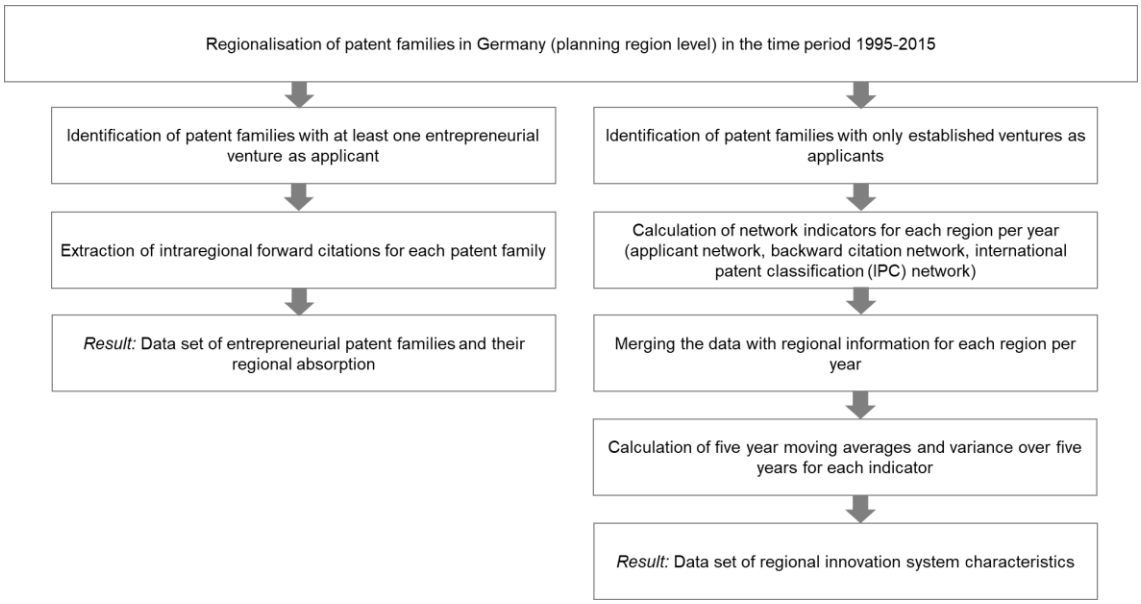


Figure 3.1.: Data set creation
Source: Own illustration

The data set creation is illustrated in Figure 3.1. Patent data from PATSTAT 2017b serves as basis for all steps and is combined with information from the data base Orbis from Bureau van Dijk. The data collection is based upon patent families in this study, which will be referred to as patents in the following.[2] Regional information are obtained from INKAR data base provided by the Federal Institute for Research on Building, Urban Affairs, and Spatial Development (BBSR) and merged to the data set. This information is based on planning regions, which are comparable to labour market areas in the USA (Fritsch & Wyrwich, 2018).

3.1.1. Creation of the entrepreneurial patent data set

Patents filed by at least one entrepreneurial venture, while defining entrepreneurial ventures as being not older than five years[3], are identified via the data base Orbis from

[2]Patent applications that share the same priority filings are clustered into a patent family. In this study DOCDB families are used. The date of the patent family is assessed by the earliest filing date of the priority filing in this family. Because this study aims to investigate the knowledge creation processes and does not aim to assess the economic value of the knowledge created, it is not differentiated between granted and non-granted patent applications.
[3]Although entrepreneurial ventures are often defined by being younger than ten years (Kollmann et al., 2020; Puig et al., 2014), there is evidence that the failing rates decrease with every year of

Bureau van Dijk. Subsidiaries and ventures having daughter companies are excluded from this sample ensuring the regional origin of the knowledge. The retrieved patents are looked up in PATSTAT 2017b to have accurate information about the address where the patent was filed. The address information of Orbis from Bureau van Dijk may not equal the address on the patent caused by moving of the company to another location. However, in some cases, the patents did not include information on the applicant address. Then, this missing data is complemented by address information of the corresponding venture from the data base Orbis from Bureau van Dijk.

As a next step, all forward citations in a five year time frame, beginning with the earliest publication date of the priority filing, are identified for each patent.[4] Taking the earliest publication date instead of the earliest filing date of the patents ensures that the patent and its content are made available for other actors. In case entrepreneurial knowledge is cited within the region of its origin, it is counted as an intraregional absorption. Citation by actors in other regions represent interregional knowledge absorption. Note that only citations regionalized within German regions can be considered for further analysis. This procedure results in a data set of entrepreneurial patents in Germany in the time frame from 1995 until 2010[5]. Overall the data set contains 1,344 patents owned by 5,346 firms that are either entrepreneurial or an incumbent firm that is collaborating with an entrepreneurial firm.

3.1.2. Creation of the regional innovation system characteristics set

The identified entrepreneurial patents are then subtracted from the set of regionalized patents in Germany in the time period from 1995 until 2015.[6] This leaves this data set only with patents filed by established ventures. Based on this, structural characteristics

---

the venture´s existence (Yang & Aldrich, 2017). This indicates that entrepreneurial ventures are getting more settled in a regional innovation system with each year of existence. To ensure that the entrepreneurial patents in this study can be viewed as mostly external to the regional innovation system, a stricter definition of entrepreneurial ventures, i.e. being not older than five years, is applied in this study.

[4]The five year time frame was chosen because 66% of all intraregional citations take place in this period. A ten year time frame would cover 87% of all intraregional citations, but would lead to a significant shrinkage of the analysable data set to the time period 1995 until 2005. Moreover, this study concentrates on the short- to mid-term effects of entrepreneurial patents because otherwise the generated knowledge might be already obsolete and does not serve as a basis for regional competitive advantages. A2 in the appendix displays the cumulative distribution of citations over years.

[5] For the years 2011 until 2015 the five year time frame to identify citations would only consist out of less than five years because the data is only available until 2015.

[6]Although not further included in the entrepreneurial patent data set, it is ensured that entrepreneurial patents from 2011 until 2015 are removed from the regional innovation system characteristics set.

of each regional innovation system for each year can be assessed. Entrepreneurial patents are not considered for this step, because this study views entrepreneurial ventures as external to the regional innovation system since these ventures are not deeply grounded in the regional networks. Incumbent firms that need to identify, assimilate, and exploit the entrepreneurially generated knowledge are interpreted as established regional innovation system.

Connecting this with the concept of Cohen and Levinthal (1989), the collective absorption of knowledge generated by entrepreneurial experimentation requires three different layers in regional networks – one for identification, one for assimilation, and one for exploitation. To assess the structural characteristics of an established innovation system, these three layers are identified for each region in each year between 1995 and 2015. Their construction is summarized in Table 3.1. First, identifying new knowledge in the regional innovation system can be interpreted as a knowledge facilitating task of connected actors in the region. A network of incumbent firms with joint innovative activities in one region enables this identification as it allows information flows between incumbent firms on multiple levels. This network is operationalized with nodes corresponding to patent applicants, and edges show a joint patent application. Second, assimilation requires an efficient learning regime in the regional innovation system. A regional network of innovations and their connection to antecedents is considered essential for the assimilation. This network is operationalized with patents as nodes and the edges between them are backward citations. All backward citations from patents are considered. Third, abilities needed to exploit new knowledge highly depend on the structure of the regional knowledge base. A network with specialized knowledge and common combinations between these describes the regional knowledge base. This network is operationalized with four-digit international patent classifications (IPC) as nodes and the edges describe the co-occurrence of IPC classes in one patent family.

Table 3.1.: Overview of networks and their operationalisation
Source: Own elaboration

Network	Rationale	Nodes	Edges
Identification	Knowledge facilitating via cooperation	Incumbent firms	Joint innovative activities
		Measurement: Patent applicants	*Measurement:* Co-occurrence of patent applicants on one patent
Assimilation	Learning regime	Innovations	Former innovations
		Measurement: Patents applications	*Measurement:* Backward citations of patent applications
Exploitation	Knowledge base	Specialized knowledge blocks	Combination of knowledge blocks
		Measurement: IPC classes	*Measurement:* Co-occurrence of IPC classes on one patent

The size of these networks, the structure, and the fragmentation as well as the stability of these network characteristics over time describe the regional innovation system context. First, the size of networks gives the framework in which knowledge flows and combinations can occur. Second, the structure of networks displays the transfer of knowledge. Third, the fragmentation gives insights about the multiplicity of regional innovation systems. Fourth, the consistency of these network characteristics plays an important role for planning innovative activities within the system.

Network indicators, which are listed in the following section, are calculated for each of these networks in every year and region. To this resulting data set, regional information (which is listed in the next section) is merged. Because the regional absorption of entrepreneurial patents is studied in a five year time window, five year moving averages for all indicators of this data set are calculated. Further, the variance of the values in this time period is determined. These mean and variance values represent the second data base – the regional innovation system characteristics in Germany in the time frame from 1995 until 2015.

3.2. Variables of the data set

3.2.1. Dependent variable – Regional absorptive capacity

The dependent variable in this study is the regional absorptive capacity of entrepreneurial knowledge. To estimate this variable, the number of entrepreneurial patents and the number of intraregional citing of these patents are used. The ratio of patents that cite entrepreneurial patents within the same region and in the following five years and entrepreneurial patents of a region in a given year forms the dependent variable. This procedure results in an efficiency ratio which captures the average intraregional citations of each entrepreneurial patent. This connects to the efficiency score proposed by Zahra and George (2002).[7] Otherwise the dependent variable would be biased by the number of entrepreneurial patents a region has. Thus, the equation for the dependent variable reads as:

$$Regional\ absorptive\ capacity = \frac{Number\ of\ patents\ citing\ the\ entrepreneurial\ patents\ within\ the\ same\ region}{Number\ of\ entrepreneurial\ patents\ in\ a\ given\ year\ in\ a\ given\ region}$$

This efficiency ratio of the absorption of entrepreneurial knowledge decides about the speed at which a regional innovation system can build up competitive advantages on basis of entrepreneurial experimentation. To compare the varying absorption capacities of different regions, this study considers only the relative value of the efficiency ratio and leaves the interpretation of the absolute values aside. Therefore, the distribution of the ratios over all regions and years is applied to convert this ratio into a factor variable with five levels. The first level contains all ratios which are zero, which corresponds to no absorption of entrepreneurial patents at all. The region in this first level of absorption rates, thus, do not use entrepreneurial experimentation as a source for competitive advantage. The remaining levels two to five correspond to the quartiles of the remaining distribution. Absorption rates in the second or third level mean to lose competitiveness over time, when having equal rates of entrepreneurial knowledge production, in comparison to other regions. Absorption rates in the level four or five enable a region to build up competitive advantages more rapid than other regions based on entrepreneurial experimentation. Table 3.2. summarizes the distribution of the observations. It is visible that the fifth level of absorption rates has the highest range of absorption rates and the highest variance of rates of all levels.

[7]The authors argue that the efficiency factor of absorptive capacity is equal to the ratio of RACAP (Realised absorptive capacity) to PACAP (Potential absorptive capacity) (Zahra & George, 2002). The measure applied in this study follows this approach and interprets intraregional citations as RACAP and the sum of entrepreneurial patents as PACAP.

Table 3.2.: Description of absorption quartiles
Source: Own illustration based on distribution of the collected data

	No absorption *(Level 1)*	**Low absorption rates *(Level 2)***	**Rather low absorption rates *(Level 3)***	**Rather high absorption rates *(Level 4)***	**High absorption rates *(Level 5)***
Number of observations	874	85	85	85	85
Range of absorption rates	0.000	0.034-0.167	0.167-0.286	0.308-0.533	0.536-7.667
Variance of absorption rates	0.000	0.001	0.001	0.005	0.902

3.2.2. Independent variables – Regional innovation system characteristics

The independent variables of interest in this study are the network indicators which describe the regional innovation system characteristics. As outlined above the chosen indicators reflect the size, the structure, the fragmentation, and the stability over time of the three networks. Table 3.3. summarizes the indicators connected to each characteristic.

First, the size of the networks is assessed by the number of edges and vertices. Second, the structure is deduced by the indicators density, transitivity, and relative diameter. While density measures the connectedness in the overall network, transitivity and diameter are calculated on basis of the biggest component of the network. Following the small-world hypothesis, transitivity is applied as a measure for clustering and diameter as a measure for longest shortest distance between two vertices. However, the diameter measurement is influenced by the network size which is corrected by taking the ratio of the diameter and the number of vertices in the biggest component. Third, the fragmentation of the networks is assessed. Because all of the variables are highly dependent on the network size, the percentage of isolated vertices from all vertices of the network, the ratio of the number of components and the number of vertices[8] in the network as well as the ratio of cutpoints and components are all taken into consideration. The isolated vertices and the number of components of a network indicate if the regional innovation system can be described as one network or rather a collection of networks

[8]A component consist out of two vertices at least. Otherwise (if it is just one vertex) it is evaluated as an isolate.

and clusters. The number of cutpoints hint at the stability of the components in the network. Lastly, the variance of all of the above mentioned indicators in five year moving windows are taken into consideration because the consistency of the variables in the citation period relates to the stability of the network characteristics.

Table 3.3.: Overview of independent network variables
Source: Own elaboration

Network characteristics	**Network indicators**
Size	Count of edges Count of vertices
Structure	Density Transitivity Relative Diameter
Fragmentation	Percentage of isolates Relative number of components Relative number of cutpoints
Stability over time	Variances of all former indicators over five years

3.2.3. Control variables - Regional context variables

Additionally, regional variables are included in the analyses as controls. These indicators, which are listed in Table 3.4., show the availability and density of human capital in the regional innovation system. The data source for these variables is INKAR provided by the Federal Institute for Research on Building, Urban Affairs, and Spatial Development (BBSR). All variables included in this study are available over the whole investigated time period on planning region level. These variables are measured as follows:[9] *Regional Potential* describes the population potential of municipalities within the perimeter of 100 kilometres. It is measured by the sum of the municipality population weighted by the area divided by thousand. This measure reflects the possibility of local interactions. The number of students is assessed by the number of students at universities and universities of applied science per 1,000 inhabitants. The share of highschool leavers with matriculation standard is the share of highschool completers that are permitted to enrol in university of all school leavers. Inhabitant density is described by the sum of inhabitants divided by the area of a region. Inhabitant-employee density is the sum of inhabitants and employees that commute to the region which are not

[9]The data sources applied for the calculation of these measures by the Federal Institute for Research on Building, Urban Affairs, and Spatial Development (BBSR) can be accessed at: https://www.inkar.de/.

inhabitants divided by the area of a region. Moreover, the variance of these indicators is taken into consideration as well.

Table 3.4.: Overview of control variables
Source: Own elaboration

	Indicators
Regional variables	Regional Potential
	Number of students
	Share of highschool leavers with matriculation standard
	Inhabitant density
	Inhabitant-employee density
	Variances of all former indicators over five years

4. Methodological approach

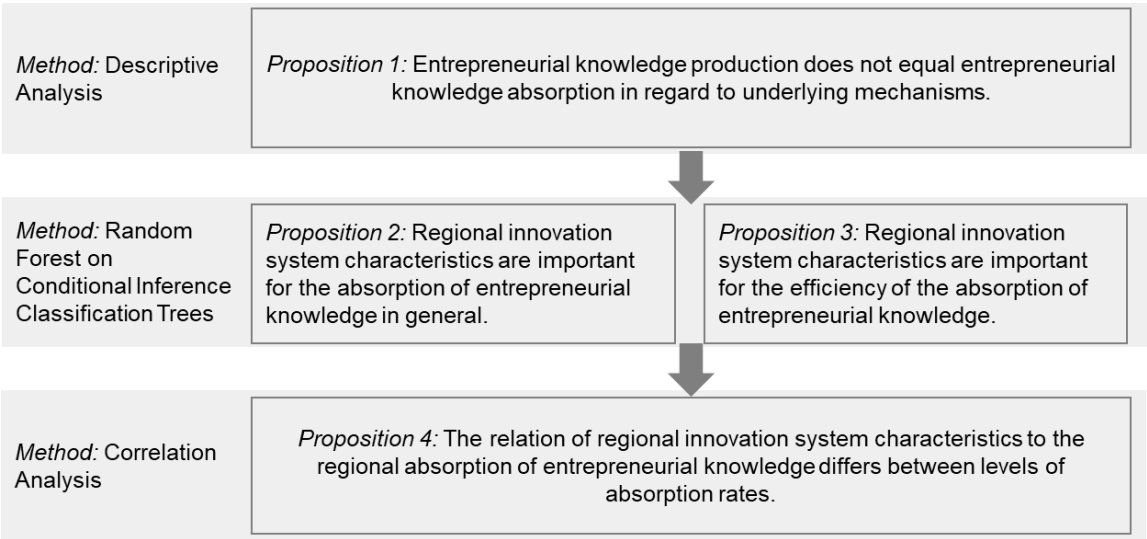


Figure 4.1.: Research design
Source: Own illustration

The research design of this study is illustrated in Figure 4.1. After *Proposition 1* is evaluated by applying descriptive analysis, *Proposition 2* and *Proposition 3* are tested by making use of random forest analysis on the basis of conditional inference classification trees. Lastly, *Proposition 4* is examined with the help of correlation analysis.

To test *Proposition 1* a descriptive analysis is performed. The geographical pattern of knowledge creation by entrepreneurs (i.e. the number of entrepreneurial patents in a region) and the knowledge absorption by incumbent firms (i.e. the ratio of entrepreneurial patents in the region and the intraregional citations of these patents) are considered in this step. To compare regional competitive advantages, not absolute but only relative

values are taken into consideration. By doing so, the following can be observed: On the one hand which regions are able to generate entrepreneurial knowledge over-proportionally, and on the other hand which regions are able to use entrepreneurial knowledge over-proportionally as source for innovative activities by incumbent firms? The first step of the analysis is thus to investigate the overlap of these to identify if regions with a high level of knowledge production show also high levels of knowledge absorption.

For the purpose to identify the most relevant regional innovation system indicators for 1) the absorption of entrepreneurial knowledge in general and 2) the efficiency of the absorption of entrepreneurial knowledge, random forest analyses based on conditional inference classification trees are conducted. There are four main reasons for this methodological approach instead of a multivariate procedure. First, medium to high correlations of the network indicators between each other poses a problem to multivariate statistical models which are most commonly used.[10] Second, the number of variables that can be imputed in a multivariate model is limited by the sample size. Third, multivariate models could be misleading in this case because these could be easily mis-specified. By taking an explorative approach, the assumption of multivariate models that all important variables are included could be violated. Fourth, multivariate models are not able to include all possible interactions between the variables considered (Strobl et al., 2009a). Conditional inference classification tree analysis overcomes these issues by following a data-driven approach which enables the use of a broad set of indicators and takes account of all possible interaction effects reflecting the complexity of regional innovation systems (Strobl et al., 2009a).

The machine learning algorithm divides the predictor space into non-overlapping regions. This means that the algorithm creates decision trees where one criterion splits the data set into two, these are further split, and so on until a stopping criterion is reached. In the case of conditional inference classification trees – an advancement of the original classification tree approaches proposed by Breiman et al. (1984) and Quinlan (1986; 1993) – the algorithm relies on the p-value of correlations between the independent variables to determine the tree's splitting criteria (Strobl et al., 2009a).[11] Using the conditional approach to build the classification trees avoids the selection bias of former algorithms that favour variables with many possible splits (Hothorn et al., 2006).

One major shortcoming of single classification trees is, however, that these are known to be sensitive to changes in the data (Strobl et al., 2009a). This is overcome by so-

[10]Considering the whole data set, the correlations between the regional innovation system indicators are higher than 0.5 in 9% of the cases. When only observations with a regional absorption of entrepreneurial knowledge higher than zero are considered, then 12% of the correlations are higher than 0.5. Full correlation matrices are available upon request.

[11]For an extensive overview of the functioning of the regression tree approaches, see Strobl et al. (2009a), and for the case of conditional inference trees, see Hothorn et al. (2006) for a detailed description.

called "ensamble" methods (Strobl et al., 2009a). Random forest, as the most prominent algorithm of this kind, creates a set of trees on random samples of the data set (Strobl et al., 2009a). To build the random samples a bootstrapping procedure is applied (Strobl et al., 2009a).[12] By comparing the multiple trees built on these bootstrapped samples, the algorithm retrieves the average importance of each variable and, thus, is a more robust measure than one single tree (Strobl et al., 2009a). Reasoned by these advantages, in this study random forest analyses on basis of conditional inference classification trees are conducted. Moreover, to avoid biased variable importance measures by high correlations between the independent variables, the conditional permutation importance is calculated.[13] This means that the covariance structure of the data set is kept equal to prevent correlated indicators to appear more important than uncorrelated ones. However, the random forest approach comes with some variability in the results by changing the starting conditions (i.e., seeds for the algorithm) and the number of trees grown. Therefore, these two conditions are varied and the results of the algorithms are compared. All variables that appear in at least two specifications with an importance score in the top ten are identified as showing robust predictive power.

To test *Proposition 2*, a random forest analysis with a binary independent variable describing if the region absorbs entrepreneurial knowledge or not (intraregional absorption of entrepreneurial knowledge equal to zero vs. intraregional absorption of entrepreneurial knowledge greater than zero) is performed. *Proposition 3* is validated by performing a separate random forest analysis with a factor variable as independent variable describing the efficiency of regional absorption of entrepreneurial knowledge (levels of knowledge absorption). This distinction into two analyses is reasoned by the expectation that the dynamics of absorbing entrepreneurial knowledge are different than the dynamics differentiating levels of absorption efficiency.

Lastly, *Proposition 4* is tested by comparing the correlation between the variables that were found to be important in the former step, and the efficiency of regional knowledge absorption of entrepreneurial knowledge. Because regional innovation systems are complex and a linear relation over all observations is unlikely, the correlation is calculated for each level of absorption rates separately by applying the Spearman correlational analysis. This procedure shows the differing importance of each variable for each level of absorption. This way the magnitude and direction of the correlations between the regional innovation system indicators and the efficiency of the regional absorption of entrepreneurial knowledge are identified.

[12]Bootstrapping is a statistical resampling method that draws multiple sub-samples of the original data set that are equal in size. This procedure is used for multiple purposes with the aim to calculate a robust result that does not depend on random variability in the original data set (James et al., 2013; Strobl et al., 2009a).
[13]For the analyses of this study, the R package "party" and the included "cforest" and conditional "varimp" functions are applied as suggested by Strobl et al. (2009b).

5. Results

5.1. Entrepreneurial knowledge production and its regional absorption

To validate *Proposition 1*, the geographical patterns of entrepreneurial knowledge production and regional absorption of entrepreneurial knowledge are compared. Figure 5.1. shows the results.

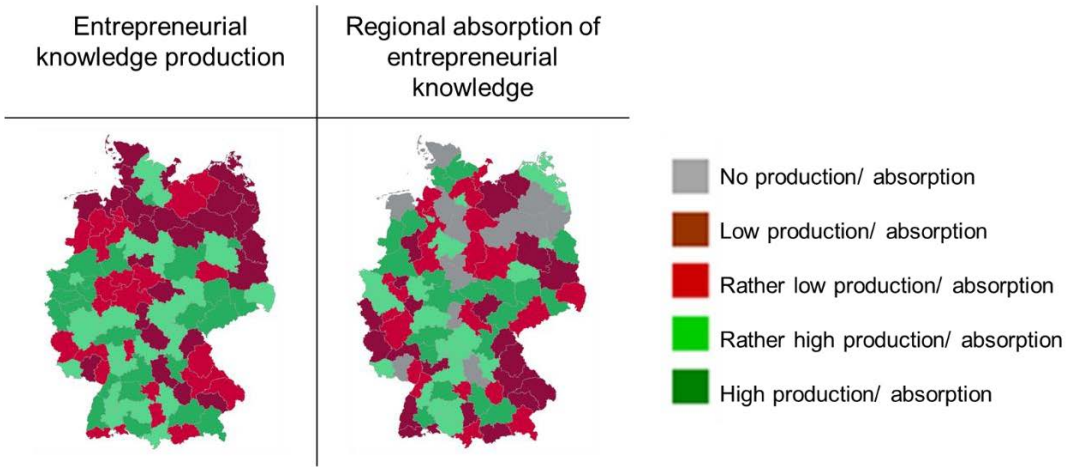


Figure 5.1.: Geographical patterns of entrepreneurial knowledge production and its regional absorption
Source: Own illustration based on calculations on basis of the collected data in the period 1995 until 2010

Looking at the entrepreneurial knowledge production, a north-south divide is visible. The northern part of Germany has more regions characterised by low and rather low entrepreneurial knowledge production than the southern part. This finding reflects that in the southern part of Germany are more big cities with more population and companies and, thus, more entrepreneurial activities. The geographical pattern of the regional absorption of entrepreneurial knowledge shows a different distribution. Although one could argue that the northern part of Germany has more regions having no absorption of entrepreneurial knowledge than the south, a clear north-south divide cannot be validated. This is a first piece of evidence that the regional knowledge production and the absorption of entrepreneurial knowledge are not equally distributed.

As a second step to complement the graphical analysis, the relevant numbers of entrepreneurial knowledge production and its regional absorption are investigated in further detail.[14] The regions with the highest entrepreneurial knowledge production, i.e. the sum of entrepreneurial patents in the period of 1995 until 2010, are Munich (910), Berlin (1101) and Stuttgart (810). This result confirms ones intuition, because these regions are metropolises and are known to be technological hubs. Turning to the regional

[14]A1 in the appendix summarises all relevant numbers of entrepreneurial knowledge production and absorption per planning region in the period 1995 until 2010.

absorption of entrepreneurial knowledge, i.e. the mean ratio of intraregional citations and entrepreneurial patents, the best performing regions are mid Schleswig-Holstein (101), western Upper Franconia (912) and northern Thuringia (1602). This comes as a surprise. Partly this result could be reasoned by the fact that these regions do have less entrepreneurial patents and it could be easier to identify, assimilate, and exploit a few entrepreneurial patents than a few hundred. However, regions with even lower entrepreneurial production rates do not show systemically higher absorption rates even the reverse is partially true since the correlation between these variables is 0.241. Finally, the correlation between the level of entrepreneurial knowledge production and the level of regional absorption is calculated, which results in a value of 0.459. Because these values can be classified as low to medium, *Proposition 1*, stating that entrepreneurial knowledge production does not equal its regional absorption, can be confirmed.

5.2. Importance of the regional innovation system characteristics for absorption

First, the networks that are considered in the following analyses are looked at in further detail. Table 5.1. summarises the descriptive statistics over all networks in the data set. The networks show great variance in the average size on all three levels – identification, assimilation, and exploitation. Moreover, the networks are on average highly clustered but show low density.

Table 5.1.: Descriptive statistics over all networks in the data set (n=1,214)
Source: Own calculation on basis of the collected data in the period 1995 until 2010

Statistic	*Mean*	*St. Dev.*	*Min*	*Pctl(25)*	*Pctl(75)*	*Max*
Identification network – Average values						
Number of edges	23.009	40.137	0	4	23.333	307.667
Number of vertices	92.695	104.883	2.667	31.708	109.167	843
Transitivity	0.429	0.246	0	0.233	0.63	0.994
Density	0.008	0.012	0	0.003	0.008	0.139
Share of isolates	0.763	0.086	0.449	0.706	0.825	1
Number of relative components	0.097	0.029	0	0.078	0.117	0.193
Number of relative cutpoints	0.108	0.106	0	0.024	0.166	0.833
Relative diameter	0.413	0.102	0	0.363	0.486	0.622
Identification network – Volatility values						
Number of edges	1,781.79	19,050.09	0	7.467	185.892	297,984.80
Number of vertices	1,063.14	5,268.42	1.2	29.792	270.192	70,130.17
Transitivity	0.142	0.093	0	0.052	0.21	0.3
Density	0.0003	0.002	0	0	0.00003	0.027
Share of isolates	0.012	0.013	0	0.004	0.016	0.116

Number of relative components	0.002	0.003	0	0.001	0.003	0.029
Number of relative cutpoints	0.022	0.057	0	0.001	0.019	0.667
Relative diameter	0.025	0.019	0	0.011	0.035	0.096
Assimilation network – Average values						
Number of edges	215.586	494.159	1	18.75	169.708	4,317.50
Number of vertices	238.935	537.932	2	27.2	197.4	4,918
Transitivity	0.074	0.061	0	0.031	0.109	0.5
Density	0.084	0.166	0.0004	0.008	0.059	1
Share of isolates	0	0	0	0	0	0
Number of relative components	0.321	0.076	0.11	0.271	0.373	0.5
Number of relative cutpoints	0.625	0.298	0	0.432	0.776	2.658
Relative diameter	0.409	0.147	0.042	0.298	0.528	0.703
Assimilation network – Volatility values						
Number of edges	11,850.66	65,364.71	0	51.167	1,959.88	1,297,516.00
Number of vertices	6,418.88	45,516.20	0	80.267	1,292.77	1,406,044.00
Transitivity	0.014	0.033	0	0.001	0.011	0.5
Density	0.015	0.043	0	0	0.0004	0.328
Share of isolates	0	0	0	0	0	0
Number of relative components	0.003	0.004	0	0.001	0.004	0.029
Number of relative cutpoints	0.075	0.2	0	0.013	0.073	2.806
Relative diameter	0.014	0.012	0	0.006	0.02	0.07
Exploitation network – Average values						
Number of edges	171.2	277.056	1	39.2	167.4	2,386
Number of vertices	107.305	79.682	3.333	46	136.292	452
Transitivity	0.549	0.162	0.167	0.428	0.659	1
Density	0.033	0.031	0.011	0.018	0.036	0.424
Share of isolates	0.212	0.074	0.051	0.167	0.244	0.852
Number of relative components	0.159	0.071	0.005	0.105	0.213	0.33
Number of relative cutpoints	1.502	1.963	0	0.451	1.785	20.5
Relative diameter	0.264	0.135	0.018	0.155	0.373	0.624
Exploitation network – Volatility values						
Number of edges	3,358.70	13,337.73	0.7	174.625	1,376.78	176,416.80
Number of vertices	254.234	450.571	2.967	51.792	209.367	3,782.27
Transitivity	0.023	0.047	0	0.003	0.02	0.267
Density	0.001	0.012	0	0.00001	0.0001	0.21
Share of isolates	0.007	0.017	0.00003	0.001	0.006	0.257

Number of relative components	0.002	0.003	0	0.0005	0.002	0.052
Number of relative cutpoints	1.881	12.278	0	0.052	0.372	181.175
Relative diameter	0.009	0.009	0	0.002	0.012	0.063
Regional variables – Average values						
Regional potential	401.669	367.845	65	200.9	438.9	2,501
Number of students	22.861	15.038	0	8.992	33.763	91.133
Share of highschool leavers with matriculation standard	26.216	6.446	14.183	21.508	30.346	55.35
Inhabitant density	368.359	537.467	44.5	138.375	363.458	3,845.00
Inhabitant-employee density	496.031	732.216	57.5	181.333	462	5,204.50
Regional variables – Volatility values						
Regional potential	19.487	153.287	0	0.7	9.067	3,905.37
Number of students	5.769	14.872	0	0.528	6.339	253.022
Share of highschool leavers with matriculation standard	15.716	22.882	0.011	0.975	25.597	152.246
Inhabitant density	39.864	280.363	0	0.8	9.367	6,187.20
Inhabitant-employee density	113.103	824.623	0	1.467	20.167	20,527.10

As a next step, the results of the random forest analyses based upon conditional inference classification trees are illustrated in Table 5.2. The variables that were identified to be important in at least two out of four specifications are marked with an “X” in the table.[15]

Reviewing the results of the analyses for the identification network, which is defined by all patent applicants in the region and their joint patenting activities, it becomes clear that the size and the fragmentation of this network influences regional absorption. To regionally absorb entrepreneurial knowledge at all, the amount of cooperation between incumbent firms in the region and the volatility of the number of clusters are most important. These findings are coherent as both indicators make the point that the structure of cooperation and the stability of cooperation clusters matters. Regarding the efficiency of regional absorption of entrepreneurial knowledge only the number of incumbent firms in the region is relevant. This result provides evidence for the importance of the symbiotic relationship of incumbent firms and entrepreneurs for a functioning regional innovation system.

The results for the assimilation network, characterised by patent applications in the region and their regional antecedents, indicate that the structure and fragmentation of the network influence absorption rates. For regional absorption in general, the volatility of the number of cutpoints, which represent bottlenecks in regional technology development trajectories, matters the most. For the efficiency of regional absorption

[15]For the results of each specification see A3 and A4 in the appendix.

processes the diameter, representing the longest trajectory of regional knowledge accumulation, is a central aspect. Thus, the regional absorption of entrepreneurial knowledge relates to the flexibility of the regional learning regime, but for an efficient absorption the construction of the main development trajectory is important.

The exploitation network, built by the knowledge blocks patents rely on and their co-occurrence, affects the regional absorption of entrepreneurial knowledge by its size, its structure, and its fragmentation. For the purpose of regional absorption in general, the number of regionally available knowledge blocks and the volatility of the share of isolated ones are most important. The exploitation network is the most important one for the efficiency of regional absorption processes as in this network the majority of important variables are located. The number of knowledge blocks available in the region and their interconnectedness remain relevant when looking at the efficiency of regional knowledge absorption. Additionally, to the volatility of isolated knowledge blocks, the average and the volatility of network density matter.

Moreover, regional conditions describing the supply in human capital, represented by regional potential and number of students, as well as their interconnectedness, represented by inhabitant and inhabitant-employee density, are relevant for regional absorption processes. The control variables *region* and *year* influence only whether there exists absorption or not, but not the efficiency of absorption processes.

Based on the empirical evidence presented, *Proposition 2* and *Proposition 3*, which state that regional innovation system indicators influence the absorption processes, can be confirmed.

Table 5.2.: Importance of variables of entrepreneurial knowledge
Source: Own elaboration based on random forest analyses

			No absorption vs. absorption		**Efficiency of absorption**	
Network	**Network characteristics**	**Network indicators**	*Mean Value*	*Variance Value*	*Mean Value*	*Variance Value*
Identification	Size	Count of edges			X	
		Count of vertices	X			
	Structure	Density				
		Transitivity				
		Relative Diameter				
	Fragmentation	Percentage of isolates				
		Relative number of components		X		
		Relative number of cutpoints				
Assimilation	Size	Count of edges				
		Count of vertices				
	Structure	Density				
		Transitivity				
		Relative Diameter				X
	Fragmentation	Percentage of isolates				
		Relative number of components				
		Relative number of cutpoints		X		
Exploitation	Size	Count of edges				
		Count of vertices	X		X	
	Structure	Density			X	X
		Transitivity				
		Relative Diameter				
	Fragmentation	Percentage of isolates		X		X
		Relative number of components				
		Relative number of cutpoints				
Control Variables		Regional Potential		X	X	
		Number of students	X			
		Share of highschool leavers with matriculation standard				
		Inhabitant density				X
		Inhabitant-employee density			X	
		Year	X			
		Region (NUTS1)	X			

5.3. Development of important variables across absorption levels

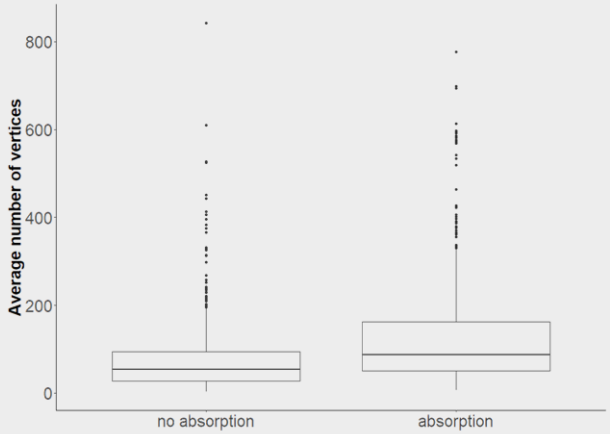
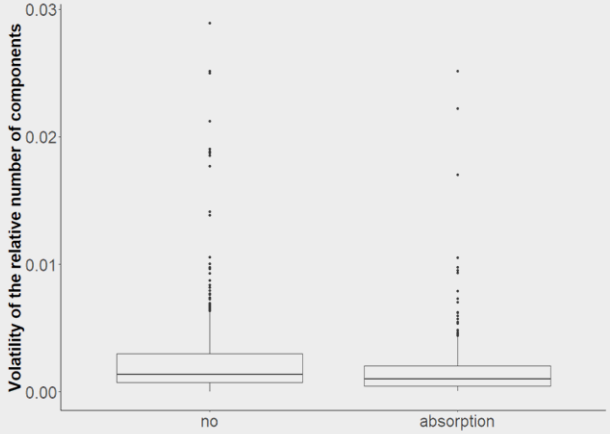
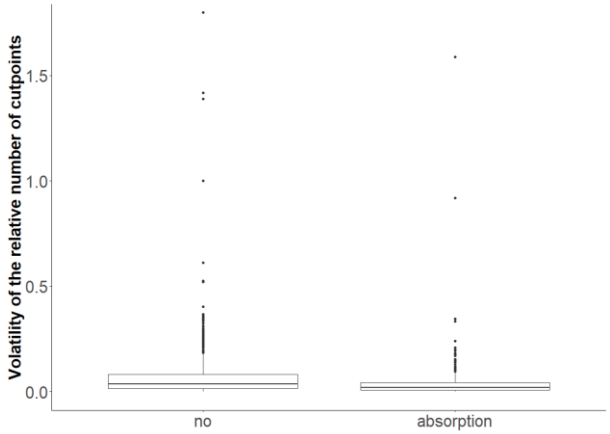
5.3.1. Regional absorption vs. no regional absorption of entrepreneurial knowledge

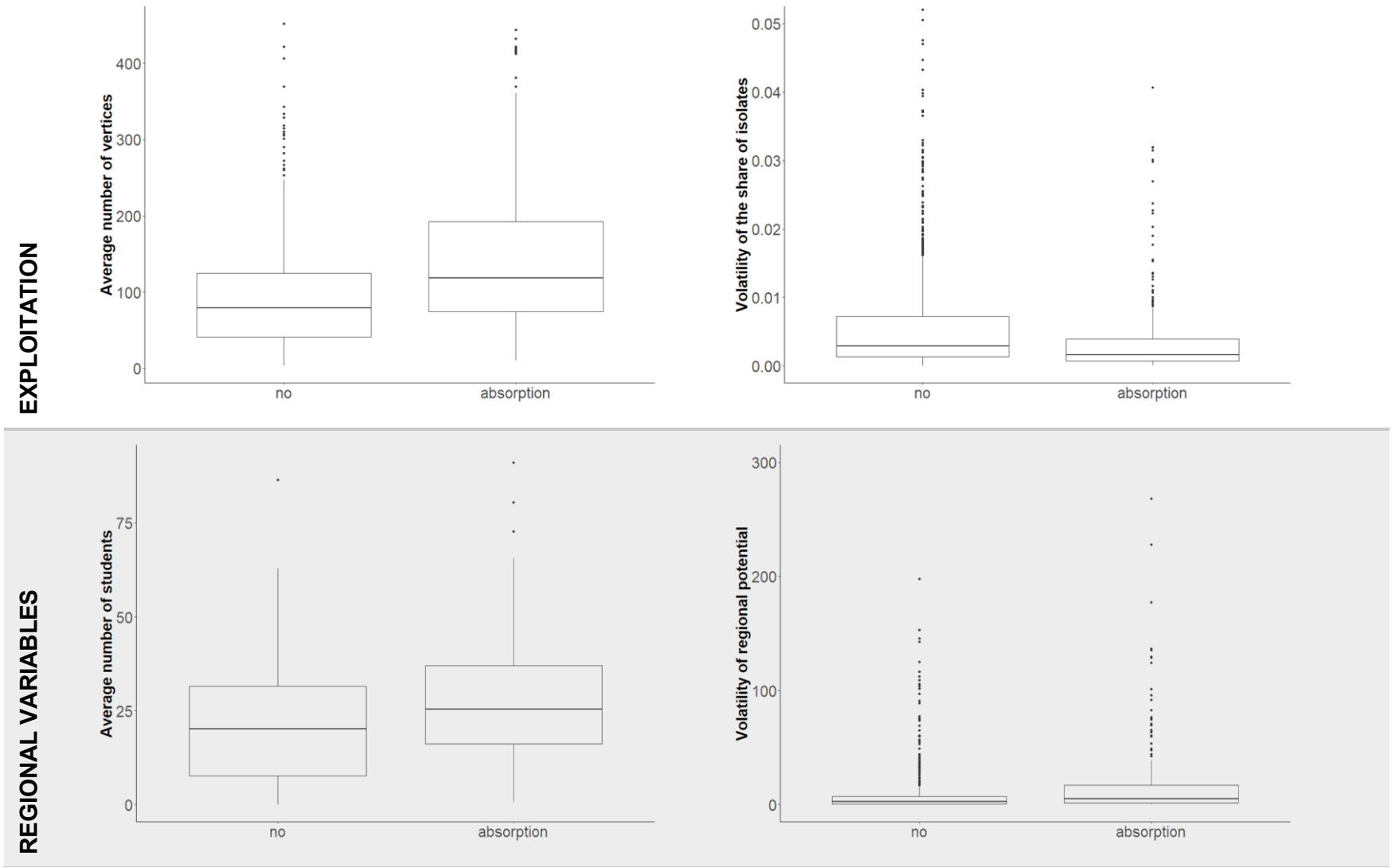
The variables that were identified as important for regional absorption of entrepreneurial knowledge in general are analysed in further detail in the following. Table 5.3. summarises the distribution of the variables important to classify a regional innovation system as absorbing or non-absorbing.

In the case of no absorption, the identification and exploitation network consist out of less vertices. This means that non-absorbing systems have less incumbent firms and less knowledge blocks available. Furthermore, these systems have less students, i.e. less human capital supply.

Additionally, regional innovation systems that do not absorb entrepreneurial knowledge are more unstable in regard to three variables: 1) the amount of components in the identification network, i.e. the amount of cooperation clusters, 2) the amount of cutpoints in the assimilation network, i.e. the amount of bottlenecks in the regional trajectories, and 3) the share of isolates in the exploitation network, i.e. unconnected knowledge blocks. These findings hint at the relevance of the predictability of developments in regional innovation systems for the absorption of entrepreneurial knowledge. However, stability in regional potential, which describes the contact potential of inhabitants, is not associated with intraregional absorption of entrepreneurial knowledge.

Table 5.3.: Comparison of important variables between absorbing and non-absorbing regional innovation systems
Source: Own illustration based upon descriptive analyses on the collected data; Note: The plots of volatility of the relative number of cutpoints in the assimilation network, volatility of the share of isolates in the exploitation network, and the volatility of regional potential omitted several outliers from the plots for the purpose of illustration

	Average values	Volatility values
IDENTIFICATION		
ASSIMILATION		


EXPLOITATION
Average number of vertices
0
100
200
300
400
no
absorption
Volatility of the share of isolates
0.00
0.01
0.02
0.03
0.04
0.05
no
absorption
REGIONAL VARIABLES
Average number of students
0
25
50
75
no
absorption
Volatility of regional potential
0
100
200
300
no
absorption

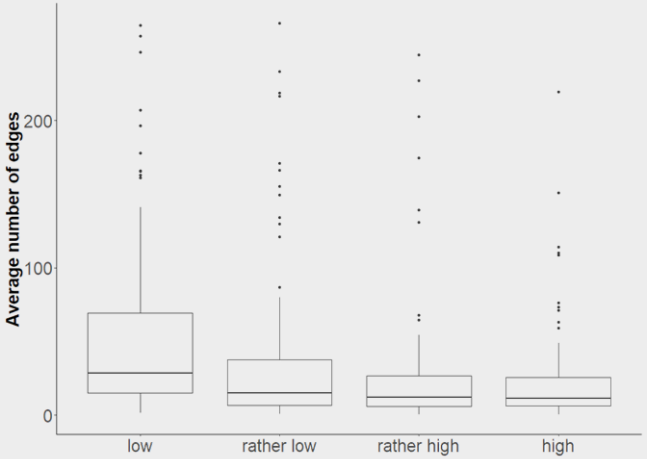
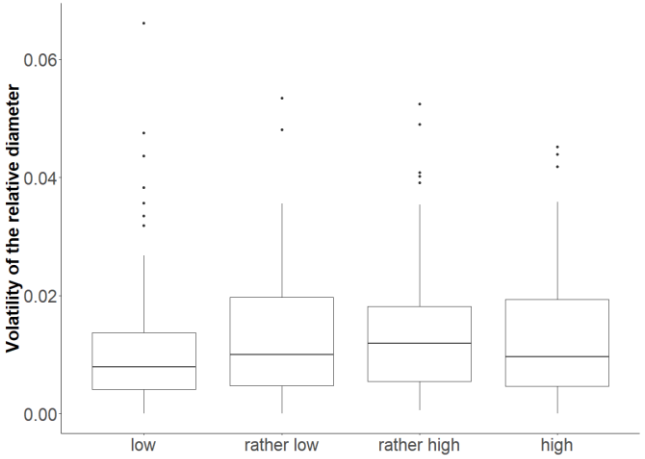
5.3.2. Efficiency of absorption

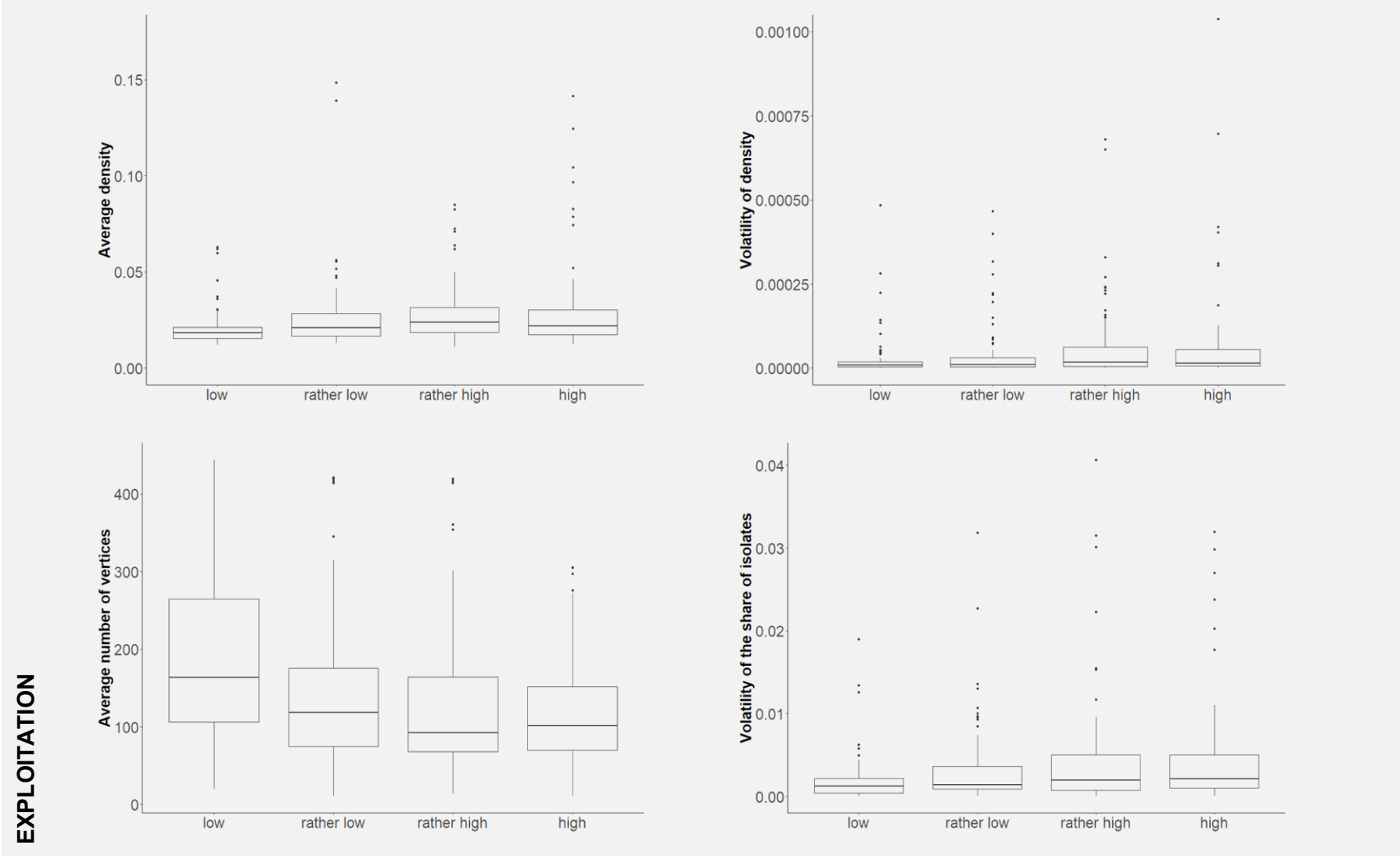
The development of the important variables across the efficiency levels of absorption of entrepreneurial knowledge is illustrated in Table 5.4.

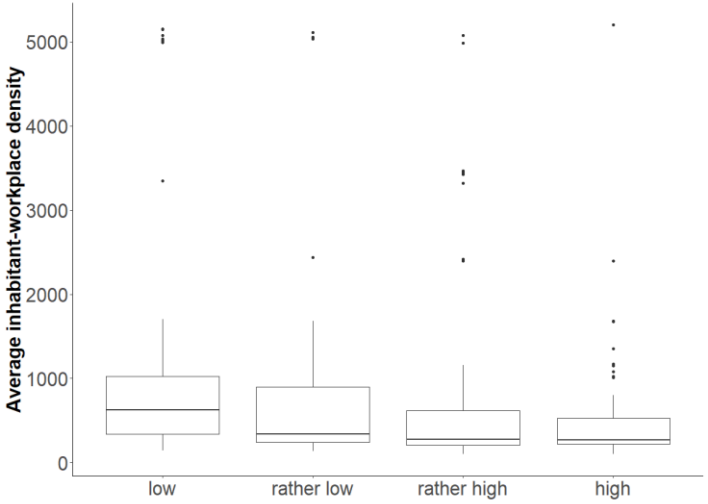
The efficiency of knowledge absorption increases with the decrease in two network indicators: 1) edges in the identification network, i.e. cooperation ties between incumbent firms and, 2) vertices in the exploitation network, i.e. regionally available knowledge blocks. Only density in the exploitation network, i.e. the connectedness of the regional knowledge base, is observed to increase over the absorption levels. These findings hint at the relevance of specialized incumbent firms in a regional innovation system. Less cooperation ties between these might imply lower possibilities for lock-ins. Further, a decreasing regional potential and inhabitant-employee density are associated with higher absorption rates. This finding might correspond to the fact that metropoles, where a high density of human capital is usual, do not rely on entrepreneurial knowledge to build a competitive advantage.

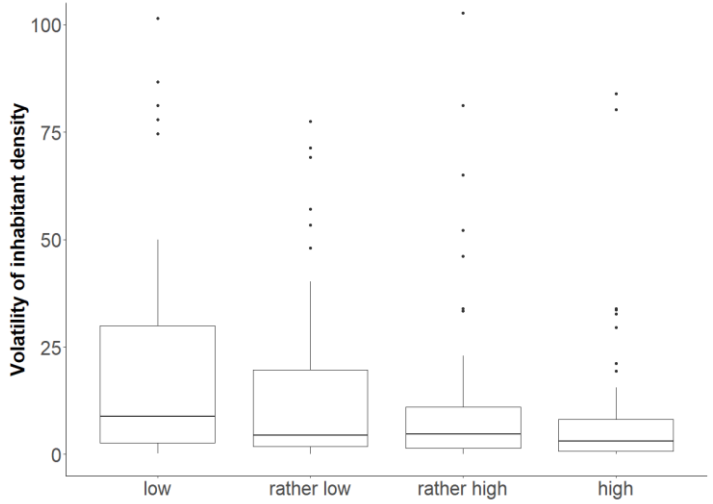
Regarding the important variables describing the volatility of network characteristics, higher knowledge absorption is associated with the increase of instability in three variables: 1) relative diameter in the assimilation network, i.e. the length of the longest trajectory, 2) share of isolates in the exploitation network, i.e. share of unconnected knowledge blocks in the regional knowledge base, and, to a lesser degree, 3) density of the exploitation network, i.e. connectedness of the regional knowledge base. These findings highlight the importance of flexible main trajectories and knowledge bases for the intraregional absorption of entrepreneurial knowledge. In contrast to this, stable rather than flexible inhabitant density of a region is associated with higher absorption levels.

Table 5.4.: Comparison of important variables between regional innovation systems with different levels of absorption
Source: Own illustration based upon descriptive analyses on the collected data; Note: The plots of average values and volatility values of the density in the exploitation network and, the volatility of inhabitant density omitted several outliers from the plots for the purpose of illustration

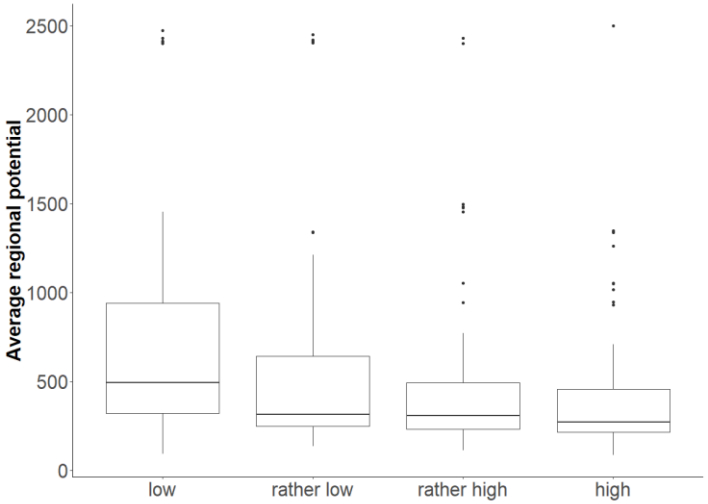
	Average values	Volatility values
IDENTIFICATION		
ASSIMILATION		


EXPLOITATION
Average density
0.15
0.10
0.05
0.00
low
rather low
rather high
high
Volatility of density
0.00100
0.00075
0.00050
0.00025
0.00000
Average number of vertices
400
300
200
100
0
Volatility of the share of isolates
0.04
0.03
0.02
0.01
0.00





REGIONAL VARIABLES



5.4. Relation between regional innovation system characteristics within levels of absorption

The former analyses revealed the variables that influence whether a region has no, low, rather low, rather high, or high absorption of entrepreneurial knowledge. In the next step, the different levels of absorption are analysed in more detail. It will be examined, which variables impact the efficiency of knowledge absorption within each absorption level. These findings help to gain a more detailed understanding of the factors that relate to success within these levels and eventually enable regions to climb one level. Table 5.5. shows all significant Spearman correlations and their respective p-value.[1]

Within the level of *low* absorption, the density of the exploitation network is positively correlated with absorption rates. A negative relation is found between the volatility of the size of the identification network and absorption rates. The volatility of the number of cutpoints in the assimilation network, i.e. the flexibility of bottlenecks in regional trajectories, is positively associated with absorption of entrepreneurial knowledge in this level. In contrast to this, population density in various specifications (regional potential, inhabitant density, and inhabitant-employee density) is negatively correlated with absorption rates. Thus, a stable sized regional innovation system, a well-connected knowledge base, and a flexible learning regime are advantages to do better than other regions in this level of knowledge absorption.

Regional innovation systems with a *rather low* absorption level show a negative relation to regional variables and their volatility. In this level, the number of students and the share of highschool leavers with matriculation standard, i.e. the supply of human capital, as well as population density, measured as inhabitant and inhabitant-employee density, are negatively related to absorption rates. Furthermore, regions with a higher flexibility in inhabitant-employee density show less absorption of entrepreneurial knowledge. Therefore, in this level, a regional innovation system having settled regional environments absorbs more entrepreneurial knowledge.

The level of *rather high* intraregional absorption of entrepreneurial knowledge shows a negative relation between volatility of transitivity in the identification network and absorption rates. Stability in the clustering of the incumbent firm network, thus, predicts higher absorption rates.

[1]The Pearson and Spearman correlations, their p-value and, in case of the Pearson correlation, the 95% confidence interval for all variables in each quartile can be found in the Tables 8.2 to Table 8.6. In the Table 5.5 the Spearman correlations are shown because there is reason to suspect that the Pearson correlations are biased because of outliers in the data. Further, regional innovation systems are complex in their nature and the detection of monotonic relations is more suitable than to expect to find linear relationships between the variables.

The highest level of absorption shows the most significant relations. Smaller networks in all three definitions – identification, assimilation, and exploitation – perform better than bigger networks. Moreover, the density of all networks is positively correlated with absorption rates within this level. The amount of cutpoints is in all three networks negatively associated with the absorption of entrepreneurial knowledge. These findings provide evidence that small, specialized, highly connected regional innovation systems have an advantage in absorbing entrepreneurial knowledge. Transitivity and relative diameter in context of the exploitation network are positively related to absorption rates, whereas in the assimilation network only relative diameter shows a positive link. The share of isolates plays a positive role in the identification and exploitation networks. Further, a higher relative number of components in the assimilation and exploitation networks predict higher absorption rates. Thus, regional innovation systems that build upon specialized clusters and cross-over connections between these clusters are favorable as well as the existence of one consistent regional learning regime. Stability in the size of all three networks, in the number of cutpoints in the exploitation network, i.e. bottlenecks of technology combinations, and in the number of students is favorable for the absorption of entrepreneurial knowledge. However, flexibility in variables of structure and fragmentation, except the number of cutpoints, is positively associated with intraregional absorption. Based upon these findings it can be derived that regional innovation systems that keep their basic characteristic - clusters which are connected via short paths – benefit from remixing clusters over time. Population density measures (regional potential, inhabitant density, and inhabitant-employee density) are negatively related to absorption rates within this level. This might hint at a lower dependency on entrepreneurial knowledge for competitive advantages in regions that are characterized by a dense supply in human capital, which is the case for metropoles.

Table 5.5.: Spearman correlations with p-value in parentheses
Source: Own calculation results
Note: Only correlations with a maximum p-value of 0.1 are shown in the Table

			Quartile 1		**Quartile 2**		**Quartile 3**		**Quartile 4**	
Net-work	**Network characteristics**	**Network indicators**	*Mean Value*	*Variance Value*	*Mean Value*	*Variance Value*	*Mean Value*	*Variance Value*	*Mean Value*	*Variance Value*
Identification	Size	Count of edges		-0.206 (0.059)					-0.336 (0.002)	-0.276 (0.011)
		Count of vertices		-0.215 (0.048)					-0.393 (0.000)	-0.292 (0.007)
	Structure	Density							0.264 (0.015)	0.339 (0.002)
		Transitivity						-0.205 (0.060)		
		Relative Diameter								0.239 (0.028)
	Fragmentation	Percentage of isolates							0.197 (0.071)	0.203 (0.063)
		Relative number of components								0.278 (0.010)
		Relative number of cutpoints							-0.420 (0.000)	
Assimilation	Size	Count of edges							-0.374 (0.000)	-0.298 (0.006)
		Count of vertices							-0.391 (0.000)	-0.311 (0.004)
	Structure	Density							0.402 (0.000)	0.407 (0.000)
		Transitivity								
		Relative Diameter							0.279 (0.010)	
	Fragmentation	Percentage of isolates								

		Relative number of components					0.286 (0.008)	
		Relative number of cutpoints		0.188 (0.085)			-0.255 (0.019)	
Exploitation	Size	Count of edges					-0.418 (0.000)	-0.383 (0.000)
		Count of vertices					-0.406 (0.000)	-0.229 (0.035)
	Structure	Density	0.303 (0.005)				0.329 (0.002)	0.262 (0.015)
		Transitivity					0.354 (0.001)	0.285 (0.008)
		Relative Diameter					0.368 (0.001)	0.218 (0.046)
	Fragmentation	Percentage of isolates					0.330 (0.002)	0.255 (0.019)
		Relative number of components					0.370 (0.000)	0.361 (0.001)
		Relative number of cutpoints					-0.378 (0.000)	-0.276 (0.011)
Control Variables		Regional Potential	-0.209 (0.055)				-0.246 (0.023)	
		Number of students		-0.272 (0.012)				-0.289 (0.007)
		Share of highschool leavers with matriculation standard		-0.212 (0.052)				
		Inhabitant density	-0.234 (0.031)	-0.213 (0.050)			-0.282 (0.009)	
		Inhabitant-employee density	-0.243 (0.025)	-0.212 (0.052)	-0.214 (0.049)		-0.290 (0.007)	

6. Discussion & Conclusion

Entrepreneurial experimentation offers the chance to extend a regional knowledge base and kick off system-wide experimentation. However, the effectiveness of the entrepreneurial experimentation to give rise to regional competitive advantages is governed by the regional absorptive capacity. The current literature argues that the absorptive capacity of a region is formed by regional context factors (Lopez-Bazo & Motellon, 2018; Malerba & McKelvey, 2020; Solano et al., 2020; Szerb et al., 2019). This study contributes to this discussion by investigating the regional innovation system characteristics that influence the regional absorption of entrepreneurial knowledge. Regionalized patent data, which is combined with firm level information from German regions over the period 1995 until 2015, build the basis for the analyses. Regional innovation systems are represented by networks on three levels: 1) applicant networks to show the cooperation between actors (identification network), 2) backward citation networks as representation for learning regimes (assimilation network), and 3) technological class networks (exploitation network) to make the knowledge base visible. Forward citations from entrepreneurial patents are used as an indicator of the success of the regional absorption process. Network indicators are applied as independent and the regional absorption success as dependent variable. These variables are implemented in a random forest analysis on basis of conditional inference classification trees.

Proposition 1 of this study, stating that the stimulation of entrepreneurial knowledge production is not sufficient to kick off system-wide absorption of this knowledge, is confirmed. Regional innovation systems that are characterised by a large amount of entrepreneurial knowledge production are not the same as those regions that make the most use of this knowledge production.

Proposition 2 and *Proposition 3*, which assume that regional innovation system indicators can explain the regional absorption of entrepreneurial knowledge in general, but also its efficiency, are confirmed. Regional innovation systems that do not show any intraregional absorption of entrepreneurial knowledge tend to have less innovative incumbent firms, a smaller knowledge base, and less human capital supply. They are also less stable in all three networks – identification, assimilation, and exploitation, which emphasises the relevance of predictability of innovation system developments for regional absorption processes. Regional innovation systems with efficient absorption of entrepreneurial knowledge are characterised by specialised incumbent firms and a flexible learning regime as well as a flexible knowledge base. The knowledge base is the most important network layer to differentiate the absorption level of regional innovation systems. In this vein, this study contributes in three ways: First, the applicant network is differentiated from the learning regime, represented by the backward citation network, to direct the attention to the distinct importance of these on regional absorptive capacity. By doing so, further evidence on the importance of applicant networks in which knowledge is shared

and of the coordination mechanisms regional absorption is based on, is provided. Also, it is shown that the regional learning regime is an important level on its own (Malerba & McKelvey, 2020; Pihlajamaa, 2018). Second, the study confirms the important role of regional knowledge bases for the absorption of new knowledge (Cohen & Levinthal, 1990; Solano et al., 2020). Third, the volatility of regional innovation system characteristics is taken into consideration. The results support that stability of network structures relate to the success and efficiency of regional absorption of entrepreneurial knowledge.

Proposition 4 is concerned with the relation of regional innovation system characteristics and the varying regional absorption rates of entrepreneurial knowledge within one absorption level – no absorption, low, rather low, rather high, and high absorption. Although the importance of regional networks is widely acknowledged in the literature, the direction of the effects of network indicators is debated. This study contributes to the literature by providing insights that the direction and magnitude of relations between regional innovation system characteristics and knowledge absorption changes on different absorption levels. Regional innovation systems with low absorption profit from a stable size, a well-connected knowledge base and a flexible learning regime. Within the level of rather low absorption rates, regional environments with less changes due to human capital supply and a lower density is beneficial. The third level with rather high absorption rates favor the stable clustering of incumbent firms. High absorption rates relate to smaller sized regional innovation systems that are flexible, but keep their key characteristic – clusters which are connected via short paths.

Overall, these findings reject the view that a larger knowledge base is per se better, as the size of the knowledge base does not appear to be important across all levels but rather multiple indicators on different levels govern the processes of regional absorption of entrepreneurial knowledge (Arvanitis & Woerter, 2015; Molina-Morales et al., 2019; Russo-Spena & Di Paola, 2019).

In most cases the regional variables, describing connectedness of the regional population, are negatively correlated with the regional absorptive capacity. This is a surprising finding since density of the population is commonly evaluated as being conducive to absorptive capacity (Molina-Morales et al., 2019; Schmidt, 2005). One explanation could be that regional innovation systems, which rely strongly on regional human capital and tacit knowledge exchange, reject ideas from unusual sources, like entrepreneurial experimentation (Pihlajamaa, 2018). Further, high density of human capital is prominently given in metropoles and these may not rely on entrepreneurial experimentation to build or sustain competitive advantages, but rely on established ventures.

The main finding of this study is that different regional mechanisms are needed to unfold the entrepreneurial potential. A policy approach focussing on only the stimulation of

entrepreneurial knowledge generation delivers inconsistent outcomes because the regional absorption processes are not considered. Therefore, policy-makers are required to monitor not only the regional entrepreneurial activities but also the mechanisms needed for regional absorption of this entrepreneurially produced knowledge (Solano et al., 2020). This goes in line with the critics of the smart specialisation approach that the diffusion of knowledge is a distinct and inevitable part for its success which requires more attention (Lindholm-Dahlstrand et al., 2019; Uyarra, 2019). To design regional policies that aim at entrepreneurial discovery processes, which enable the transformation of whole regions, requires a detailed understanding of regional absorption processes. Besides that, machine learning techniques, like the random forest analysis conducted in this study, demonstrate great explanatory power for the kind of data analysed in this study Often, for the analysis of regional innovation systems, where indicators are highly dependent on each other, commonly applied multivariate statistics are unable to provide satisfying results. Therefore, machine learning techniques are of special importance for future research.

Nevertheless, the scope of the study is limited, paving the road for future research. First, regionalized patent data is used as basis for the analysis. This comes with the difficulty that the regional entrepreneurial knowledge and its regional absorption is underestimated because not every patent includes address information. However, there is no reason to assume that the subset of patents with regional information is systematically different from the patents that include information on the location. Second, patent data is a proxy for the innovative activities in the regional innovation system. Due to the non-patenting of certain industries, this data is incomplete. Future research could complement patent data with other types of data, e.g., surveys or news data, to capture a holistic picture of regional innovative activities. Third, the methodological approach in this study, conducting a random forest analysis based upon conditional inference classification trees, comes with the limitation that the used model is pooled. Due to computational restrictions it was also not possible to control for regional effects for every region (96 planning regions) but only on federal state level. Future research may use the data on a more sophisticated model allowing the exploitation of the panel structure of the data, controlling for regional effects. Fourth, the regional innovation systems, which are pooled in this analysis, are different in size. This can cause multiple network indicators to be influenced. However, as these network indicators capture unique aspects, the explanatory power of each indicator is not eliminated. As the random forest analysis based upon conditional inference trees is not influenced by correlations between explanatory variables, the procedure has the ability to retrieve the unique contribution of each variable. Fifth, the scope of the variables analysed in this study is not comprehensive but rather reflects the basic network indicators. It could be beneficial for future research to focus on one class of indicators and compare different specifications of these. Lastly, the split of the observations into the second to fifth level of absorption, which is based on the distribution of the data, comes with a limitation. As seen before, the variance in the fourth quartile, the fifth level of absorption, is greater than in the other

three. This may explain why the fourth quartile shows more relations to regional innovation system characteristics. Although, this split has the benefit that the conditional inference classification trees are not biased due to imbalances in class sizes, future research may validate the results by choosing different data splits.

References

Acs, Z., & Armington, C. (2004). Employment Growth and Entrepreneurial Activity in Cities. *Regional Studies*, 38(8), 911–927. https://doi.org/10.1080/0034340042000280938

Acs, Z. J., Braunerhjelm, P., Audretsch, D. B., & Carlsson, B. (2009). The knowledge spillover theory of entrepreneurship. *Small Business Economics*, 32(1), 15–30. https://doi.org/10.1007/s11187-008-9157-3

Arvanitis, S., & Woerter, M. (2015). Exploration or Exploitation of Knowledge from Universities: Does It Make a Difference? *Economics of Innovation and New Technology*, 24(5–6), 596–623.

Asheim, B. T., & Gertler, M. S. (2006). The Geography of Innovation: Regional Innovation Systems. *Oxford University Press*. https://doi.org/10.1093/oxfordhb/9780199286805.003.0011

Asheim, B. T., & Isaksen, A. (2002). Regional innovation systems: the integration of local 'sticky'and global 'ubiquitous' knowledge. *The Journal of Technology Transfer*, 27(1), 77-86.

Balland, P.-A., & Rigby, D. (2017). The Geography of Complex Knowledge. *Economic Geography*, 93(1), 1–23. https://doi.org/10.1080/00130095.2016.1205947

Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1), 31–56. https://doi.org/10.1191/0309132504ph469oa

Baumol, W. J. (2002). Entrepreneurship, innovation and growth: The David-Goliath symbiosis. *Journal of Entrepreneurial Finance*, JEF, 7(2), 1-10.

Boschma, R. (2017). Relatedness as driver of regional diversification: A research agenda. *Regional Studies*, 51(3), 351–364. https://doi.org/10.1080/00343404.2016.1254767

Boschma, R., Heimeriks, G., & Balland, P.-A. (2014). Scientific knowledge dynamics and relatedness in biotech cities. *Research Policy*, 43(1), 107–114. https://doi.org/10.1016/j.respol.2013.07.009

Braunerhjelm, P., Acs, Z. J., Audretsch, D. B., & Carlsson, B. (2010). The missing link: Knowledge diffusion and entrepreneurship in endogenous growth. *Small Business Economics*, 34(2), 105–125. https://doi.org/10.1007/s11187-009-9235-1

Breimann, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and regression trees. *Pacific Grove,* Wadsworth.

Camisón, C., & Forés, B. (2010). Knowledge absorptive capacity: New insights for its conceptualization and measurement. *Journal of Business Research*, 63(7), 707–715. https://doi.org/10.1016/j.jbusres.2009.04.022

Cantner, U., & Graf, H. (2008). Interaction structures in local innovation systems. *Jena Economic Research Papers*, No.2008, 040.

Caragliu, A., & Nijkamp, P. (2012). The impact of regional absorptive capacity on spatial knowledge spillovers: The Cohen and Levinthal model revisited. *Applied Economics*, 44(11), 1363–1374. https://doi.org/10.1080/00036846.2010.539549

Carlsson, B., Braunerhjelm, P., McKelvey, M., Olofsson, C., Persson, L., & Ylinenpää, H. (2013). The evolving domain of entrepreneurship research. *Small Business Economics*, 41(4), 913–930. https://doi.org/10.1007/s11187-013-9503-y

Carpenter, R. E., & Petersen, B. C. (2002). Capital market imperfections, high-tech investment, and new equity financing. *The Economic Journal*, 112(477), F54-F72.

Choi, J.-D., & Park, J.-H. (2017). The Performance Effect of Two Different Dimensions of Absorptive Capacity and Moderating Role of Holding-Cash. *Technology Analysis and Strategic Management*, 29(9), 1033–1047.

Cohen, W. M., & Levinthal, D. A. (1989). Innovation and Learning: The Two Faces of R & D. *The Economic Journal*, 99(397), 569. https://doi.org/10.2307/2233763

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

Colombo, M. G., & Grilli, L. (2007). Funding Gaps? Access To Bank Loans By High-Tech Start-Ups. *Small Business Economics*, 29(1–2), 25–46. https://doi.org/10.1007/s11187-005-4067-0

D'Adda, D., Guzzini, E., Iacobucci, D., & Palloni, R. (2019). Is Smart Specialisation Strategy Coherent with Regional Innovative Capabilities? *Regional Studies*, 53(7), 1004–1016.

David, P. A., & Foray, D. (2002). An introduction to the economy of the knowledge society. *International Social Science Journal*, 54(171), 9–23. https://doi.org/10.1111/1468-2451.00355

Distel, A. P. (2019). Unveiling the Microfoundations of Absorptive Capacity: A Study of Coleman's Bathtub Model. *Journal of Management*, 45(5), 2014–2044. https://doi.org/10.1177/0149206317741963

Fernandez-Serrano, J., Martinez-Roman, J. A., & Romero, I. (2019). The Entrepreneur in the Regional Innovation System. A Comparative Study for High- and Low-Income Regions. *Entrepreneurship and Regional Development*, 31(5–6), 337–356.

Ferreira Moutinho, R. F. (2016). Absorptive Capacity and Business Model Innovation as Rapid Development Strategies for Regional Growth. *Investigacion Economica*, 75(295), 157–202.

Foray, D. (2018). Smart specialization strategies as a case of mission-oriented policy—A case study on the emergence of new policy practices. *Industrial and Corporate Change*, 27(5), 817–832. https://doi.org/10.1093/icc/dty030

Foray, D., David, P. A., & Hall, B. H. (2011). Smart specialization. From academic idea to political instrument, the surprising career of a concept and the difficulties involved in its implementation. *Management of Technology & Entrepreneurship*, No.2011-001.

Foray, D., Goddard, J., Beldarrain, X. G., Landabaso, M., McCann, P., Morgan, K., Nauwelaers, C., & Ortega-Argilés, R. (2012). Guide to Research and Innovation Strategies for Smart Specialisations (RIS 3). *European Comission*.

Fritsch, M. (2008). How does new business formation affect regional development? Introduction to the special issue. *Small Business Economics*, 30(1), 1–14. https://doi.org/10.1007/s11187-007-9057-y

Fritsch, M., & Kublina, S. (2018). Related variety, unrelated variety and regional growth: The role of absorptive capacity and entrepreneurship. *Regional Studies*, 52(10), 1360–1371. https://doi.org/10.1080/00343404.2017.1388914

Fritsch, M., & Wyrwich, M. (2018). Regional knowledge, entrepreneurial culture, and innovative start-ups over time and space—an empirical investigation. *Small Business Economics*, 51(2), 337–353. https://doi.org/10.1007/s11187-018-0016-6

Furman, J. L., & Hayes, R. (2004). Catching up or standing still?: National innovative productivity among 'follower'countries, 1978–1999. *Research policy*, 33(9), 1329-1354.

Grillitsch, M. (2019). Following or Breaking Regional Development Paths: On the Role and Capability of the Innovative Entrepreneur. *Regional Studies*, 53(5), 681–691.

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased Recursive Partitioning: A Conditional Inference Framework. *Journal of Computational and Graphical Statistics*, 15(3), 651–674. https://doi.org/10.1198/106186006X133933

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning (Vol. 103). New York. Springer. https://doi.org/10.1007/978-1-4614-7138-7

Kollmann, T., Jung, P.B., Kleine-Stegemann, L., Ataee, J., & de Cruppe, K. (2020). Deutscher Startup Monitor 2020. Innovation statt Krise.

Lane, P. J., Koka, B., & Pathak, S. (2002). A Thematic Analysis and Critical Assessment of Absorptive Capacity Research. In *Academy of Management proceedings* (Vol. 2002, No. 1, pp. M1-M6). Briarcliff Manor, NY 10510: Academy of Management.

Lindholm-Dahlstrand, A., Andersson, M., & Carlsson, B. (2019). Entrepreneurial Experimentation: A Key Function in Systems of Innovation. *Small Business Economics*, 53(3), 591–610.

Lopez-Bazo, E., & Motellon, E. (2018). Innovation, Heterogeneous Firms and the Region: Evidence from Spain. *Regional Studies*, 52(5), 673–687.

Malerba, F., & McKelvey, M. (2020). Knowledge-Intensive Innovative Entrepreneurship Integrating Schumpeter, Evolutionary Economics, and Innovation Systems. *Small Business Economics*, 54(2), 503–522.

Mason, C., & Brown, R. (2014). Entrepreneurial ecosystems and growth oriented entrepreneurship. *Final report to OECD*, Paris, 30(1), 77-102.

Molina-Morales, F. X., Martinez-Chafer, L., & Valiente-Bordanova, D. (2019). Disruptive Technology Adoption, Particularities of Clustered Firms. *Entrepreneurship and Regional Development*, 31(1–2), 62–81.

Neffke, F., Hartog, M., Boschma, R., & Henning, M. (2018). Agents of Structural Change: The Role of Firms and Entrepreneurs in Regional Diversification. *Economic Geography*, 94(1), 23–48. https://doi.org/10.1080/00130095.2017.1391691

Pihlajamaa, M. (2018). Absorbing Radical Ideas from Unusual Sources—The Role of Social Integration Mechanisms. *Technology Analysis and Strategic Management*, 30(2), 131–143.

Puig, F., González-Loureiro, M., & Ghauri, P. N. (2014). Internationalisation for survival: The case of new ventures. *Management International Review*, *54*(5), 653-673.

Pyka, A., Kudic, M., & Muller, M. (2019). Systemic Interventions in Regional Innovation Systems: Entrepreneurship, Knowledge Accumulation and Regional Innovation. *Regional Studies*, 53(9), 1321–1332.

Qian, H., & Jung, H. (2017). Solving the Knowledge Filter Puzzle: Absorptive Capacity, Entrepreneurship and Regional Development. *Small Business Economics*, 48(1), 99–114.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106. https://doi.org/10.1007/BF00116251

Quinlan, JR. (1993). C4.5: Programms for Machine Learning. *Morgan Kaufmann Publishers Inc*, San Francisco.

Rothaermel, F. T., & Deeds, D. L. (2004). Exploration and exploitation alliances in biotechnology: A system of new product development. *Strategic Management Journal*, 25(3), 201–221. https://doi.org/10.1002/smj.376

Russo-Spena, T., & Di Paola, N. (2019). Inbound Open Innovation in Biopharmaceutical Firms: Unpacking the Role of Absorptive Capacity. *Technology Analysis and Strategic Management*, 31(1), 111–124.

Schmidt, T. (2005). What determines absorptive capacity. In *DRUID summer conference*.

Schumpeter, J.A. (1934). The Theory of Economic Development: An Inquiry into Profits, Capital, Credits, Interest, and the Business Cycle. *Transaction Publishers*, Piscataway.

Solano, G., Larraneta, B., & Aguilar, R. (2020). Absorptive Capacity Balance and New Venture Performance: Cultivating Knowledge from Regional Clusters. *Technology Analysis and Strategic Management*, 32(11), 1264–1276.

Spigel, B. (2017). The Relational Organization of Entrepreneurial Ecosystems. *Entrepreneurship Theory and Practice*, 41(1), 49–72. https://doi.org/10.1111/etap.12167

Storper, M., & Venables, A. J. (2004). Buzz: Face-to-face contact and the urban economy. *Journal of Economic Geography*, 4(4), 351–370. https://doi.org/10.1093/jnlecg/lbh027

Strobl, C., Hothorn, T., & Zeileis, A. (2009b). Party on! A New, Conditional Variable Importance Measure for Random Forests Available in the party Package. *Technical Report*, No. 50.

Strobl, C., Malley, J., & Tutz, G. (2009a). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, 14(4), 323–348. https://doi.org/10.1037/a0016973

Szerb, L., Lafuente, E., Horvath, K., & Pager, B. (2019). The Relevance of Quantity and Quality Entrepreneurship for Regional Performance: The Moderating Role of the Entrepreneurial Ecosystem. *Regional Studies*, 53(9), 1308–1320.

Teirlinck, P., & Spithoven, A. (2019). The R&D Knowledge Base in City-Agglomerations and Knowledge Searching in Product Innovative SMEs. *Entrepreneurship and Regional Development*, 31(5–6), 516–533.

Tushman, M. L., & Anderson, P. (1986). Technological Discontinuities and Organizational Environments. *Administrative Science Quarterly*, 31(3), 439. https://doi.org/10.2307/2392832

Uyarra, E. (2010). What is evolutionary about 'regional systems of innovation'? Implications for regional policy. *Journal of Evolutionary Economics*, 20(1), 115–137. https://doi.org/10.1007/s00191-009-0135-y

Uyarra, E. (2019). Smart Specialization as Place-based Policy: Lessons Learnt?. *Regional Insights*, *2019*, 1-6. https://doi.org/10.1080/13673882.2018.00001022

Yang, T., & Aldrich, H. E. (2017). "The liability of newness" revisited: Theoretical restatement and empirical testing in emergent organizations. *Social Science Research*, 63, 36–53. https://doi.org/10.1016/j.ssresearch.2016.09.006

Zahra, S. A., & George, G. (2002). Absorptive Capacity: A Review, Reconceptualization, and Extension. *Academy of Management Review*, 27, 185–203.

Appendix

A1: Regional entrepreneurship activity and regional absorptive capacity
Source: Own calculation based on calculations on basis of the collected data in the period 1995 until 2010

		Knowledge production		**Knowledge absorption**		
Planning region code (ROR Code)	*German planning region name*	*Sum of entrepreneurial patents*	*Quantile in knowledge production*	*Sum of intraregional citations on entrepreneurial patents*	*Mean ratio of entrepreneurial patents and intraregional citations*	*Quantile in knowledge absorption*
101	Schleswig-Holstein Mitte	58	3	30	0.517	4
103	Schleswig-Holstein Ost	19	1	2	0.105	2
104	Schleswig-Holstein Süd	74	3	7	0.095	2
105	Schleswig-Holstein Süd-West	15	1	3	0.200	4
201	Hamburg	173	4	28	0.162	3
301	Braunschweig	67	3	8	0.119	2
302	Bremen-Umland	32	2	3	0.094	2
303	Bremerhaven	13	1	1	0.077	2
304	Emsland	31	2	7	0.226	4
307	Hannover	79	3	14	0.177	3
309	Lüneburg	13	1	1	0.077	2
310	Oldenburg	31	2	10	0.323	4
311	Osnabrück	34	2	1	0.029	1
401	Bremen	54	2	9	0.167	3
501	Aachen	129	4	5	0.039	1
502	Amsberg	45	2	8	0.178	3
503	Bielefeld	108	4	8	0.074	2
504	Bochum/Hagen	142	4	12	0.085	2

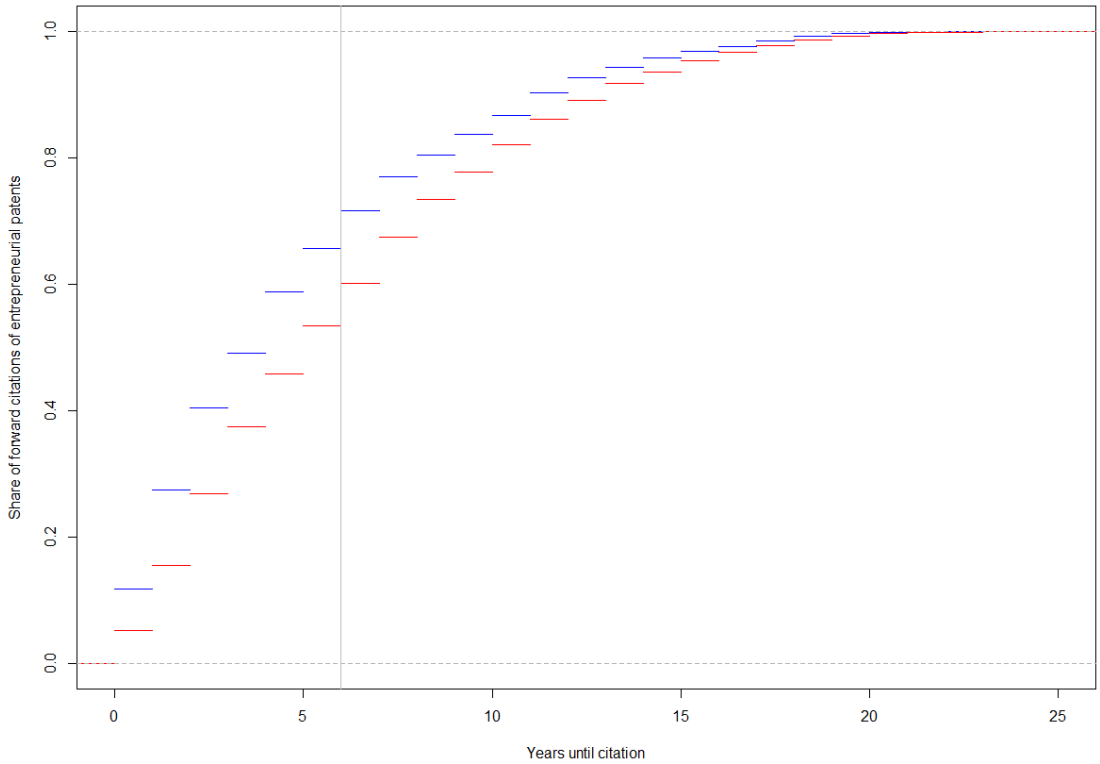
505	Bonn	61	3	5	0.082	2
506	Dortmund	97	4	4	0.041	1
507	Duisburg/Essen	155	4	26	0.168	3
508	Düsseldorf	212	4	33	0.156	3
509	Emscher-Lippe	54	3	6	0.111	2
510	Köln	137	4	24	0.175	3
511	Münster	133	4	33	0.248	4
512	Paderborn	53	2	13	0.245	4
513	Siegen	38	2	2	0.053	1
601	Mittelhessen	70	3	10	0.143	3
602	Nordhessen	32	2	6	0.188	4
604	Rhein-Main	190	4	52	0.274	4
605	Starkenburg	76	3	15	0.197	4
701	Mittelrhein-Westerwald	73	3	8	0.110	2
702	Rheinhessen-Nahe	53	2	3	0.057	1
703	Rheinpfalz	40	2	3	0.075	2
704	Trier	49	2	1	0.020	1
801	Bodensee-Oberschwaben	97	4	17	0.175	3
802	Donau-Iller (BW)	65	3	9	0.138	3
803	Heilbronn-Franken	97	4	18	0.186	3
804	Hochrhein-Bodensee	39	2	1	0.026	1
805	Mittlerer Oberrhein	105	4	12	0.114	2
806	Neckar-Alb	63	3	9	0.143	3
807	Nordschwarzwald	70	3	11	0.157	3
808	Ostwürttemberg	38	2	4	0.105	2
809	Schwarzwald-Baar-Heuberg	55	3	19	0.345	4
810	Stuttgart	347	4	69	0.199	4
811	Südlicher Oberrhein	88	4	5	0.057	1

812	Unterer Neckar	78	3	4	0.051	1
901	Allgäu	56	3	6	0.107	2
902	Augsburg	63	3	19	0.302	4
903	Bayrischer Untermain	45	2	8	0.178	3
904	Donau-Iller (BY)	50	2	2	0.040	1
905	Donau-Wald	52	2	1	0.019	1
906	Industrieregion Mittelfranken	154	4	23	0.149	3
907	Ingolstadt	20	1	2	0.100	2
908	Landshut	25	1	1	0.040	1
909	Main-Röhn	22	1	3	0.136	3
910	München	464	4	97	0.209	4
911	Oberfranken-Ost	30	1	2	0.067	1
912	Oberfranken-West	54	3	28	0.519	4
913	Oberpfalz	41	2	2	0.049	1
914	Oberpfalz-Nord	32	2	1	0.031	1
915	Regensburg	35	2	2	0.057	1
916	Südostoberbayern	82	4	11	0.134	2
918	Würzburg	58	3	10	0.172	3
1001	Saar	77	3	11	0.143	3
1101	Berlin	456	4	79	0.173	3
1201	Havelland-Fläming	77	3	4	0.052	1
1202	Lausitz-Spreewald	19	1	1	0.053	1
1203	Oderland-Spree	30	1	8	0.267	4
1302	Mittleres Mecklenburg/Rostock	39	2	1	0.026	1
1303	Vorpommern	14	1	2	0.143	3
1304	Westmecklenburg	39	2	1	0.026	1
1401	Oberes Elbtal/Osterzgebirge	192	4	36	0.188	4
1402	Oberlausitz-Niederschlesien	68	3	5	0.074	2

1403	Südsachsen	208	4	22	0.106	2
1404	Westsachsen	137	4	33	0.241	4
1502	Anhalt-Bitterfeld-Wittenberg	43	2	6	0.140	3
1503	Halle/S.	79	3	26	0.329	4
1504	Magdeburg	116	4	12	0.103	2
1601	Mittelthüringen	54	3	1	0.019	1
1602	Nordthüringen	26	1	15	0.577	4
1603	Ostthüringen	152	4	55	0.362	4
1604	Südthüringen	77	3	8	0.104	2
102	Schleswig-Holstein Nord	17	1	0	0	0
305	Göttingen	35	2	0	0	0
306	Hamburg-Umland-Süd	8	1	0	0	0
308	Hildesheim	17	1	0	0	0
312	Ost-Friesland	19	1	0	0	0
313	Südheide	25	1	0	0	0
603	Osthessen	4	1	0	0	0
705	Westpfalz	26	1	0	0	0
917	Westmittelfranken	25	1	0	0	0
1204	Prignitz-Oberhavel	22	1	0	0	0
1205	Uckermark-Barnim	3	1	0	0	0
1301	Mecklenburgische Seenplatte	8	1	0	0	0
1501	Altmark	1	1	0	0	0

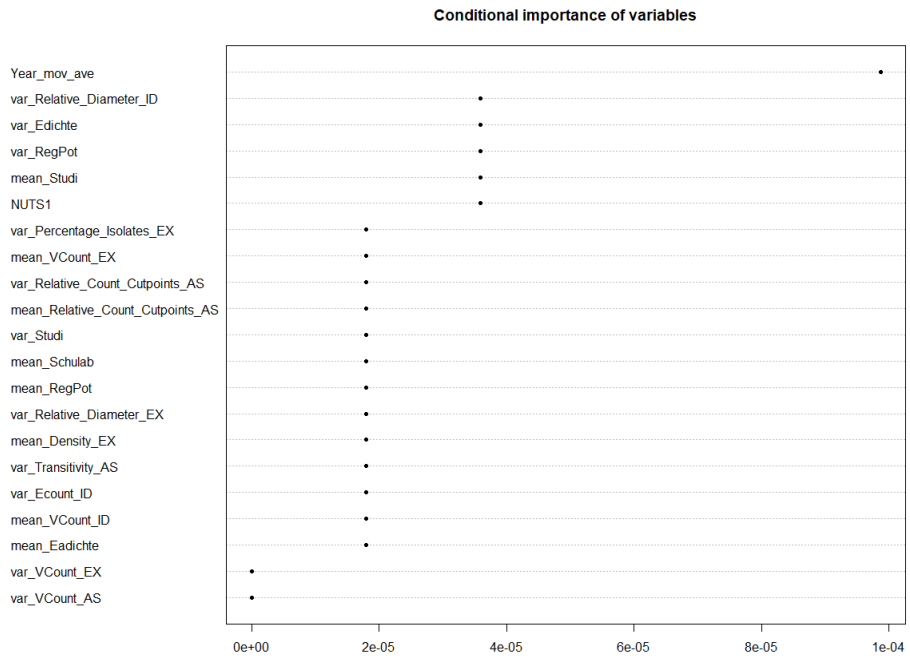
A2: Cumulative distribution of intraregional citations (blue) and interregional citations (red)

Source: Own illustration based on collected data in the period 1995 until 2010

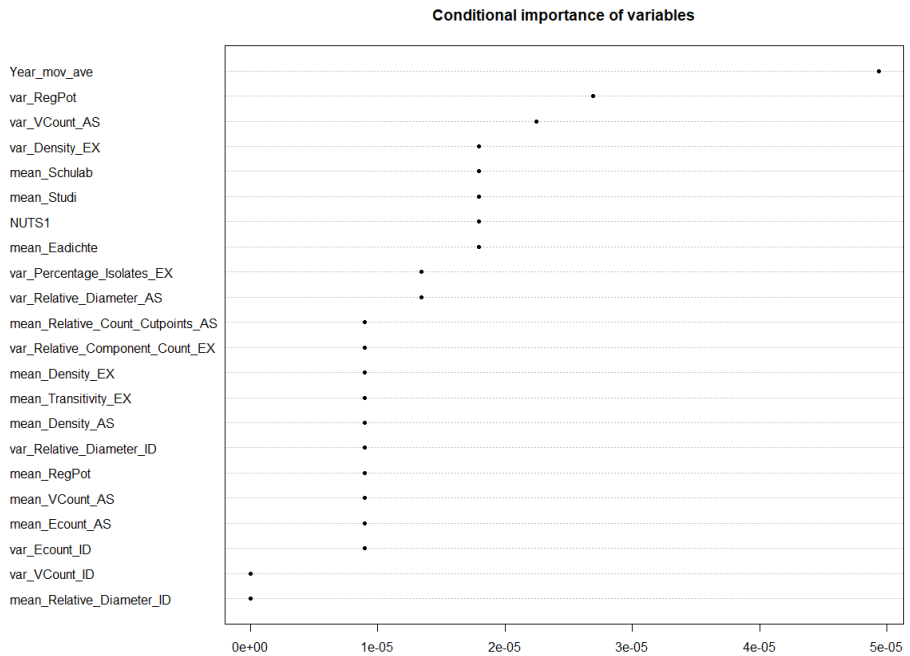


A3: Specifications of random forest analysis for no absorption vs. absorption of entrepreneurial knowledge
Source: Own illustration based on random forest analysis on collected data

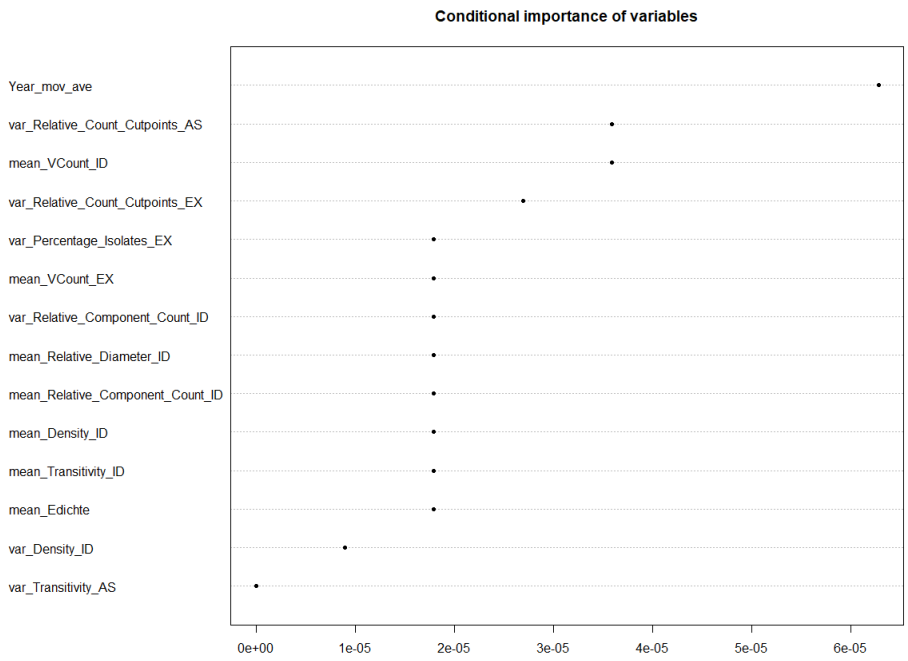
No Absorption vs. Absorption



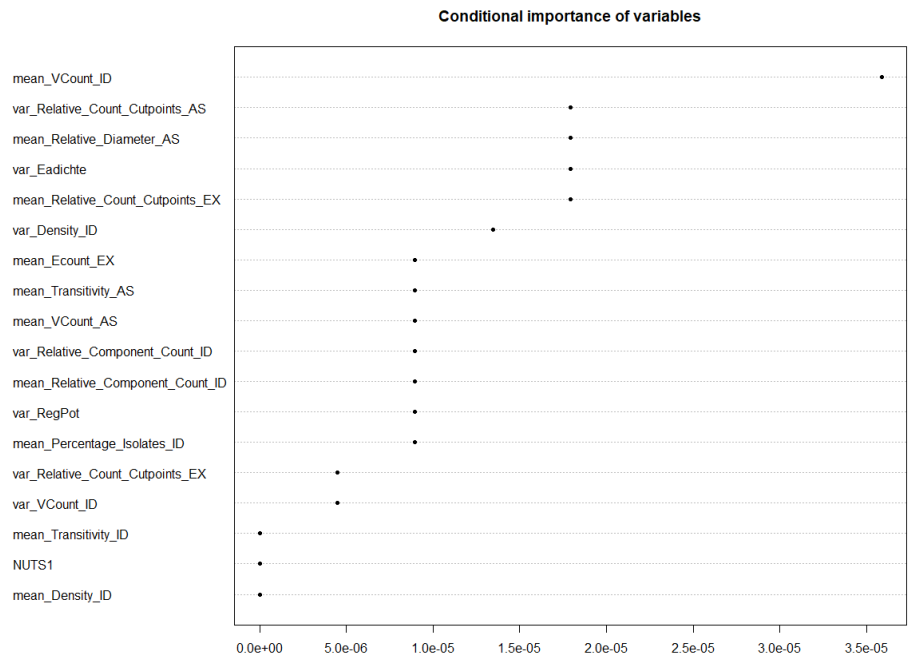
Seed: 1234
Number of trees: 250
Correct classification rate: 74.629%



Seed: 1234
Number of trees: 500
Correct classification rate: 74.794%



Seed: 2021
Number of trees: 250
Correct classification rate: 74.712%



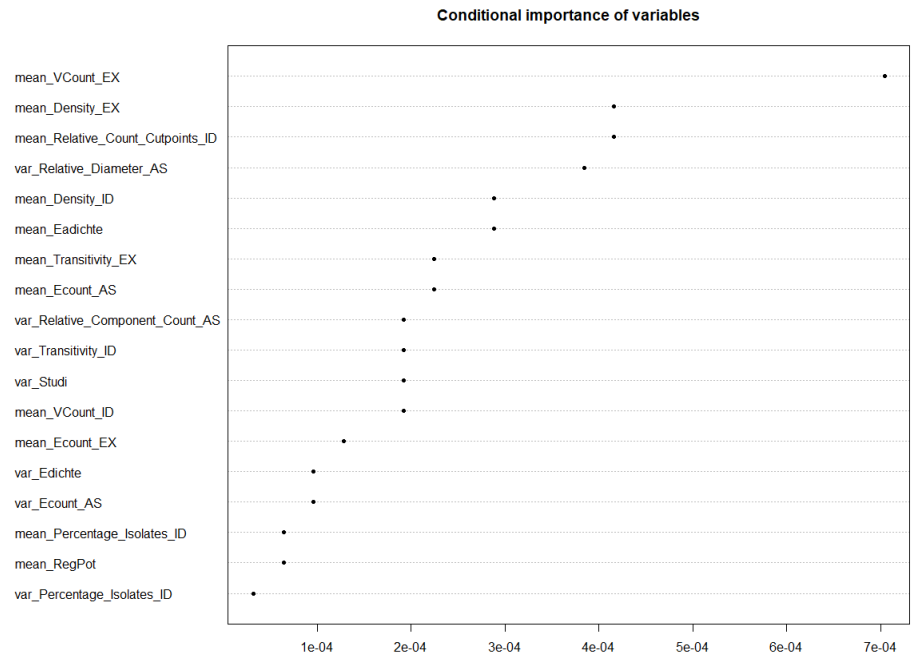
Seed: 2021
Number of trees: 500
Correct classification rate: 74.712%

Variables
(replace XX with ID for identification, AS for assimilation and EX for exploitation network)

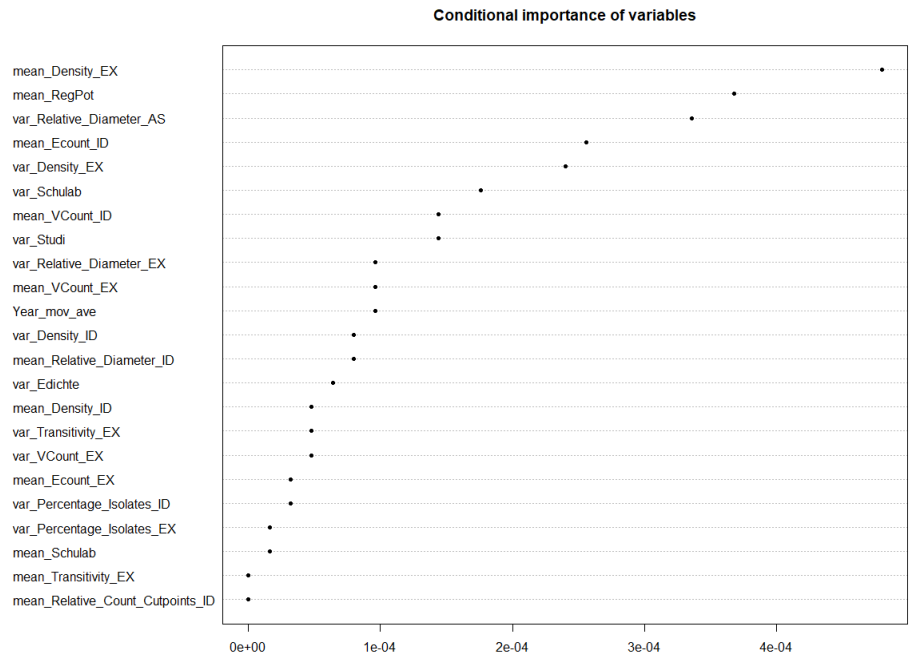
Average / Volatility regional potential	mean_RegPot / var_RegPot
Average / Volatility number of students	mean_Studi / var_Studi
Average / Volatility share of highschool leavers with matriculation standard	mean_Schulab / var_Schulab
Average / Volatility inhabitant density	mean_Edichte / var_Edichte
Average / Volatility inhabitant-employee density	mean_Eadichte / var_Eadichte
Average / Volatility count of edges	mean_Ecount_XX / var_Ecount_XX
Average / Volatility count of vertices	mean_VCount_XX / var_VCount_XX
Average / Volatility transitivity	mean_Transitivity_XX / var_Transitivity_XX
Average / Volatility density	mean_Density_XX / var_Density_XX
Average / Volatility share of isolates	mean_Percentage_Isolates_XX / var_Percentage_Isolates_XX
Average / Volatility number of relative components	mean_Relative_Component_Count_XX / var_Relative_Component_Count_XX
Average / Volatility number of relative cutpoints	mean_Relative_Count_Cutpoints_XX / var_Relative_Count_Cutpoints_XX
Average / Volatility relative diameter	mean_Relative_Diameter_XX / var_Relative_Diameter_XX

A4: Specifications of random forest analysis for efficiency of regional absorption of entrepreneurial knowledge
Source: Own illustration based on random forest analysis on collected data

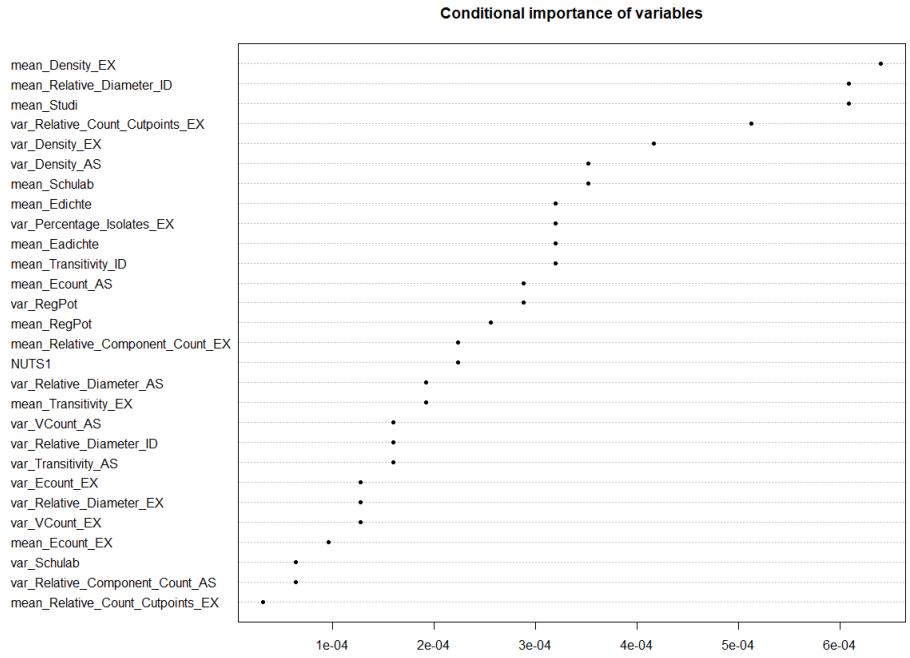
Efficiency of absorption



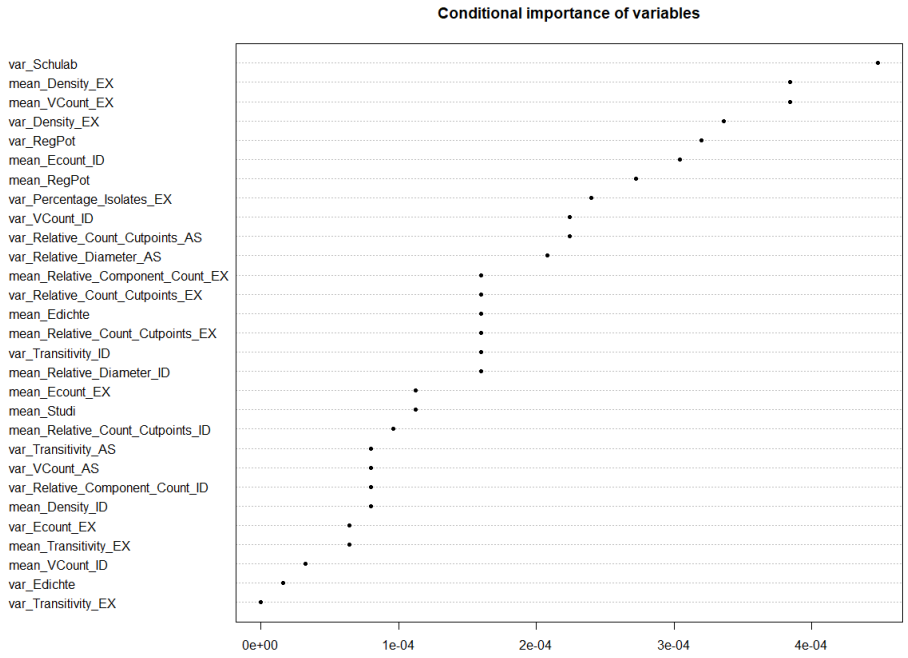
Seed: 1234
Number of trees: 250
Correct classification rate: 28.235%



Seed: 1234
Number of trees: 500
Correct classification rate: 30.294%



Seed: 2021
Number of trees: 250
Correct classification rate: 27.941%



Seed: 2021
Number of trees: 500
Correct classification rate: 28.824%

Variables
(replace XX with ID for identification, AS for assimilation and EX for exploitation network)

Average / Volatility regional potential	mean_RegPot / var_RegPot
Average / Volatility number of students	mean_Studi / var_Studi
Average / Volatility share of highschool leavers with matriculation standard	mean_Schulab / var_Schulab
Average / Volatility inhabitant density	mean_Edichte / var_Edichte
Average / Volatility inhabitant-employee density	mean_Eadichte / var_Eadichte
Average / Volatility count of edges	mean_Ecount_XX / var_Ecount_XX
Average / Volatility count of vertices	mean_VCount_XX / var_VCount_XX
Average / Volatility transitivity	mean_Transitivity_XX / var_Transitivity_XX
Average / Volatility density	mean_Density_XX / var_Density_XX
Average / Volatility share of isolates	mean_Percentage_Isolates_XX / var_Percentage_Isolates_XX
Average / Volatility number of relative components	mean_Relative_Component_Count_XX / var_Relative_Component_Count_XX
Average / Volatility number of relative cutpoints	mean_Relative_Count_Cutpoints_XX / var_Relative_Count_Cutpoints_XX
Average / Volatility relative diameter	mean_Relative_Diameter_XX / var_Relative_Diameter_XX

Imprint

Bremen Papers on Economics & Innovation

Published by
University of Bremen, Faculty of Business Studies & Economics,
Institute for Economic Research and Policy (ierp)
Max-von-Laue-Straße 1, 28359 Bremen, Germany

Bremen Papers on Economics & Innovation #2205

Responsible Editor: Prof. Dr. Jutta Günther

ISSN 2629-3994