Guiding discovery

How knowledge development and innovation strategies are directed by the regional context

Master's Thesis | Master Innovation Sciences
Copernicus Institute, Utrecht University
16th of August 2020

Author:
Tim Willemse BSc.
t.p.h.willemse@students.uu.nl
Student number: 4257349

Supervisor:
dr. Gaston Heimeriks

Second reader:
dr. ir. Matthijs Janssen

Abstract

Regional policy makers struggle in designing effective bottom-up and mission-oriented innovation policies that are embedded in their territorial context. Knowledge development has proven to be fundamental in these innovation dynamics and is said to be evolving in relation to the regional knowledge capabilities and institutional context. Therefore, this study aims to find how specific attributes of the regional knowledge base and institutional context relate to knowledge development in European regions and how these dynamics shape a region's prioritisation strategy in the RIS3 program of the European Union. This is done from a scientific knowledge perspective, as this provides insight in the fundamental regional capabilities from which economic and societal goals can be addressed. A series of quantitative analyses showed that the related diversification opportunities provided by the regional knowledge base, quality of government and institutional thickness have a positive relation to complex knowledge development. These results confirm the expectation that both the regional knowledge capabilities and the institutional context are instrumental in knowledge development dynamics. In the RIS3 program, a thematic approach might be advised for socially relevant topics, since the overrepresentation of priorities in green technology and health might hamper the alignment with the regional context. However, it was shown that in general regions are able to prioritise according to knowledge base capabilities in terms of strategy complexity and relatedness. Furthermore, it was found that the regional representation of the higher education sector positively influences strategy complexity and that the ability of the government to connect to regional actors enhances the strategy relatedness. These results did not show that regions currently utilise the option to integrate diversity in the prioritisation strategy, while this could be a viable option, to improve future diversification capabilities, especially for lagging regions. Lastly, as expected, advanced regions were found to possess the most capabilities in the subjected institutional features related to prioritisation. However, intermediate regions were found to possess promising institutional capabilities as well and even the lagging regions showed a few institutional features which might provide some perspective in future smart specialisation efforts. To conclude, by recognizing the fundamental role that knowledge and institutional elements play in both knowledge development and innovation policy, more effectively designed territorial innovation strategies can be developed.

Acknowledgments

First of all, I want to thank my supervisors dr. Gaston Heimeriks and dr. ir. Matthijs Janssen for guiding me through the research process and providing me with valuable feedback. Additionally, I would like to thank the Centre for Science and Technology Studies (CWTS) and Leiden University for their hospitality and for helping me in my data collection. Specifically, I would like to thank Sonia Mena Jara, dr. Ingeborg Meijer and Josephine Bergmans for the inspiration and feedback they gave for my thesis and their pleasant cooperation in the European Commission's CHERRIES project, to which this study shares common ground. Furthermore, my gratitude goes to Alfredo Yegros for his help in the data collection and operationalisation process. My gratitude also goes to Mihály Fazekas for his instructions on obtaining the relevant indicators from his research. Lastly, I want to thank my fellow students for making it a memorable and instructive journey in the Master program. Especially, I would like to thank Emma Boersma, Shamiram Abdulahad and Freek Boelders for their support during the thesis process.

Table of content

1. Introduction

Historically, regional innovation policy was often implemented top-down as part of a national agenda, trying to stimulate economic and technical development and reducing inter regional disparities (Rodriquez-Pose, 2013). From the 1980s onwards, regions started to focus on facilitating bottom-up movements such as stimulating entrepreneurial activity, upgrading technological capacity and improving the regional skill base (Lagendijk & Cornford, 2000). Recently, this approach is complemented, in the light of societal challenges like climate change, with policies focusing on a system's perspective and taking a mission-oriented standpoint (McCann & Ortega-Argilés, 2015; Mazzucato, 2011). To accommodate this bottom-up, mission-oriented political environment, the European Union (EU) has launched its arguably most ambitious regional policy strategy till date with the Road to Investment Strategy 3 (RIS3) program based on the smart specialisation rationale (Morgan, 2017). The European Commission (2012) sets three main goals to be achieved in the program, that are smart, sustainable and inclusive growth. With the smart specialisation concept, regions are instructed to realise these goals by selecting priority investment fields for specialisation. Regions can do this smartly, by basing them on existing regional capabilities. The priorities can be targeted towards one or multiple sectors, technologies, societal challenges and cultural and natural resources (Gianelle et al., 2018). The resulting collection of priority fields ultimately composes a region's priority portfolio (Gianelle et al., 2018; McCann & Ortega-Argilés, 2015).

Practically, the priority fields for each EU region are set by a process of entrepreneurial discovery, in which regional actors use their regional background to discover innovation opportunities (Foray et al., 2011; OECD, 2013). This unique bottom-up approach tries to create alignment between regional capabilities and regional policy, which the smart specialisation concept is looking for. Additionally, it enables regions to prioritise new domains that the actors view as important (Foray, 2016). However, policy makers struggle with the practical implication of a prioritisation strategy. They find it hard to switch from centralised policies to policy making based on regional actor input and regional context (Landabaso, 2014). In this regard, prioritising the right fields and regional elements without neglecting others is seen as a challenge (Landabaso & Foray, 2014). Subsequently, they often opt for broad topics such as stopping climate change and improving healthcare, without a clear rationale and consequently reducing the overall effectiveness of the smart specialisation policy (Trippl et al., 2019; Kroll, 2015). Clearly, there is a strong social need to understand the relationship between a region's contextual features and innovation dynamics to support policy makers in the prioritisation process (Boschma, 2017).

Since Schumpeter's (1943) seminal work, knowledge has taken a central role within the innovation policy arena, as it proved its role in facilitating economic and technological gains. In more recent times, the notion rose that the importance of knowledge has only increased in the modern economy, which is therefore often labelled as the knowledge-based economy (Sahal, 1981; Cooke & Leydesdorff, 2006). This is caused by a shift from a manufacturing- to a service-based economy and by products becoming more complex, which both enhance knowledge requirements (Powell & Snellman, 2004). Moreover, knowledge is viewed as a key driver in facing current societal challenges due to their intricacy (Shiroyama et al., 2012; van der Hel, 2016). Hence, the importance that knowledge development holds in the smart specialisation process and its fundamental role for achieving the innovation that the RIS3 program pursues (Boschma & Gianelle, 2014).

To successfully implement knowledge development in regional policy strategies, research has indicated that first of all the nature of a region's knowledge base has to be considered in the smart specialisation process (Boschma & Gianelle, 2014; Carayannis & Grigoroudis, 2016; Balland et al., 2019). The regional knowledge base is the aggregate of accumulated knowledge and experience of organisations in the region (Cantner et al., 2010). Regions tend to build upon the accumulated knowledge and experience in order to develop new knowledge. Given that knowledge is unevenly distributed over regions, regions are faced with varying levels of complexity, diversity and specialisation as departing points in their strategies (Florida, 2005; Heimeriks et al., 2019). Acknowledging these circumstances, it is recommended that regions select their priority fields according to the relatedness of their knowledge base towards new topics (Balland et al., 2019; Cebolla & Navas, 2019; Heimeriks & Boschma, 2013). However, it is unknown if regions are capable of implementing these principles and how a region's scientific portfolio influences their prioritisation (Heimeriks & Balland, 2015).

Secondly, there is a consensus in the scientific community that the institutional context plays an essential interactive role in knowledge and smart specialisation dynamics (Boschma, 2004; Rodriquez-Pose, 2013; Sotarauta, 2018; Foray et al., 2018; Benner, 2019). Rodriquez-Pose (2013) defines the institutional context as a combination of formal (e.g. rules and laws) and informal (e.g. relationships and norms) institutions. Therefore, the institutional context is different for different types of regions, based on different forms and levels of human capital, industrial structures, institutional systems and governance capabilities, shaped by (inter)national and local logics (McCann & Ortega-Argilés, 2013, 2014). These regional attributes support and constrain the environment in which knowledge production and entrepreneurship take place (Boschma, 2004). They are sticky in space and not easily changed, making the institutional context a persistent regional characteristic. Therefore, each region has a unique institutional arena in which knowledge development unfolds and policy strategies are formed, that subsequently could affect the manner in which these processes are carried out (Rodriquez-Pose, 2013). However, it is yet unclear how exactly the institutional context and knowledge development relate and how it guides the selection of priority fields. Despite recent qualitative efforts in understanding the interplay between the institutional context and smart specialisation, it remains highly contested which institutional factors are important and unclear how important these factors are (Trippl et al., 2019; Benner, 2019).

Vedula et al. (2019) found, looking at the aforementioned regional elements, that a supportive institutional logic has a larger effect on firm entrance in the United States when a knowledge pool is larger and a smaller effect when the knowledge pool is more specialised. Regarding the RIS3 program, Benner (2019) found that institutional dynamics in the entrepreneurial discovery led to governmental learning as an important by-product of smart specialisation. Trippl et al. (2019) confirm that the regional institutional context is a factor for prioritisation and found that RIS3 leads to policy reorientation and system transformation in advanced regions and to policy learning and system building attempts in less developed regions. However, these studies fail to address how the knowledge-institutional configuration behaves in terms of knowledge development and how regions strategise their prioritisation accordingly (Gianelle et al., 2018). To address this, and previous raised gaps, the following research question is formulated:

How do the regional knowledge base and institutional context relate to knowledge development and how does this influence the prioritisation strategy in European regions?

This study attempts to answer this question in quantitative manner, making use of Web of Science publication data, scientific and public data on regional characteristics and the Eye@RIS3 database on regional RIS3 priorities. There is opted for publication data, because this provides insight on fundamental knowledge capabilities and it also provides information on non-technical and not (yet) commercially exploited knowledge development (Heimeriks et al., 2019).

Theoretically this contributes to determining which institutional attributes play a role in regional knowledge development. Moreover, combining this with knowledge base characteristics provides a more explicit and holistic understanding on regional knowledge development. On top of that, it provides understanding on how these regional knowledge development dynamics affect the design of the most recent innovation strategies. Combining these two key findings, this study can ultimately assess how regional contexts, knowledge development dynamics and innovation policy interact. For these purposes, novel methodological approaches are introduced in this study to map a region's institutional context and assess its prioritisation strategy. From a societal perspective, this study enhances the ability of policy makers to effectively align their regional capabilities with the design of regional development targets. Additionally, for the European Union, it provides insight on how different types of regions should be instructed to perform their prioritisation in the smart specialisation program.

First, the dynamics and interacting attributes of regional knowledge development, knowledge base capabilities, the institutional context and RIS3 prioritisation from previous literature are theoretically described in section two. Along this discussion, hypotheses are formed on the relations between these concepts. In the methodological section three, meaningful regional indicators are constructed that are qualitatively assessed and then subjected to a regression and a cluster analysis. Further reflection on the results from these analyses presented in section four is described in section five. Lastly, conclusions are drawn in section six.

2. Theory

In this section, first the relevance of knowledge in smart specialisation is explained from an evolutionary perspective. Secondly, the relevance of the institutional context from this perspective is displayed and expectations on the relations with knowledge development are deductively derived. Lastly, the prioritisation process is described in more detail to assess the possible effects of the region's context on its prioritisation strategy.

2.1. Regional knowledge development

Generally, a distinction is made between fundamental and applicable knowledge in the context of knowledge development (Adams, 1990). Fundamental knowledge development aims to acquire basic understanding, often pursued in the form of scientific research (Nelson & Romer, 1996; Cardinal et al., 2001). Applicable knowledge development builds on this basis with the purpose of generating knowledge useful for solving societal needs, often in the form of patents or trademarks (Adams, 1990). Since smart specialisation aims at achieving societal visions, both fundamental and applicable knowledge are important to achieve these goals. Fundamental knowledge indicates which underlying capabilities are generated and applicable knowledge shows on which practical implementations actors are working (Adams, 1990; Nelson & Romer, 1996; Cardinal et al., 2001). Given that this research primarily focuses on the relations between underlying regional capabilities and prioritisation, a fundamental scientific knowledge perspective is taken, when looking at knowledge development. Additionally, fundamental knowledge production is more easily influenced by public interventions than other forms of knowledge development, making it an interesting focus for smart specialisation policy (Nelson, 1959). With this perspective on knowledge development the reasoning of previous literature on this topic is followed (Boschma et al., 2014; Heimeriks & Balland, 2015; Heimeriks et al., 2019).

This literature stream assesses knowledge development as both path and place dependent. Knowledge development is place dependent because of its tacit character, embodied in networks and routines, that are difficult to imitate and sticky in space (Maskel & Malmberg 1999). Due to this spatial concentration, knowledge development trajectories often take place on a regional level (Frame et al., 1977; Heimeriks & Boschma, 2014). This makes knowledge the main source for a region's unique competitive advantage and provides smart specialisation with its regional rationale. (Balland et al., 2019). Knowledge is path dependent because it accumulates, building on existing knowledge and providing opportunity for further knowledge development (Arthur, 2009). This is important, because technological trajectory theory argues that technologies that develop along a technological trajectory have more chance of succeeding than technologies that do not, due to the acquired knowledge that accumulates along the trajectory (Dosi, 1982; Dosi & Nelson, 2013). Along this trajectory, previous 'simpler topics' are likely to be forgotten and would require a new learning process to acquire again (Arthur, 2009). Consequently, it is expected that regions concentrate on a limited number of scientific fields and can be characterised by distinct scientific portfolios (Boschma et al., 2014; Heimeriks et al., 2019).

Applying these concepts in the regional context, created the notion that regional development targets would prove more fruitful when building upon existing local capabilities (Lambooy & Boschma, 2001). This culminated in the understanding that regions can specialise smartly by diversifying their knowledge base into new fields by building on related local capabilities (Boschma & Gianelle, 2014). Here, related(ness) is defined as the degree of which similar sets of cognitive capabilities and skills are required. Therefore, regions that produce knowledge closely related to new topics, have the best capabilities to keep up with evolving patterns at the forefront of markets and society (Balland et al., 2019). Furthermore, regions that produce unrelated knowledge to new topics would find it difficult to follow these trends and find themselves locked-in in previous paradigms (Heimeriks & Balland, 2014).

To limit the risks and uncertainties incurred by unrelated diversification, the diversity of knowledge present in the region is important. Diversity can be defined as the degree of variety, balance and disparity between scientific fields (Stirling, 2007). The higher the diversity, the more likely it is that new topics are related to the regional knowledge base (Heimeriks et al., 2019). Diversity could therefore ease the process of engaging in new knowledge development by increasing the amount of related diversification opportunities. By providing these opportunities, knowledge base diversity makes regions more flexible in responding to unexpected changes in its social or market environment and subsequently can help regions to prevent ending up in a locked-in state (Morgan, 2017). On top of that, diversity is required to pursue complex knowledge trajectories (Hidalgo & Hausmann, 2009). Hidalgo & Hausmann (2009) argue that the complexity of a territorial entity or technological field is indicated by the diversity of knowledge produced and its ubiquity. Regions want to diversify into complex fields, because these fields require sophisticated knowledge, which is highly path and place dependent, providing long lasting competitive advantage (Maskel & Malmberg, 1999, Hidalgo & Hausmann 2009). Knowledge base diversity is important in this respect, because complex fields rely on a diverse set of knowledge capabilities by definition and it provides the related diversification opportunities needed to sustain the long trajectories towards complex fields.

Given that knowledge development trajectories and subsequent branching opportunities are strongly geographically bounded, regions are expected to differ greatly in terms of knowledge base diversity and complexity (Hidalgo & Hausmann, 2009). This also implies that regions are bounded by their knowledge base characteristics, regarding the amount of related diversification opportunities they are realistically presented with (Balland et al., 2019). Consequently, it is expected that the regions with more diversification opportunities have a better chance of selecting complex trajectories and can therefore generally achieve more complex knowledge development. Additionally, the spiky nature of the knowledge distribution across regions would imply that only a limited number of regions possess the diversification capabilities to achieve high knowledge complexity (Heimeriks et al., 2019). Presumably, this influences the prioritisation options for the majority of regions included in the smart specialisation program that do not enjoy the capabilities reflected by a highly capable knowledge pool. For these regions producing complex knowledge is likely a risky endeavour that requires a longer learning process.

H1: Related diversification opportunities are positively related to complex knowledge development

2.2. Institutional context

The evolutionary knowledge perspective does not paint a complete picture on knowledge development and regional innovation on its own. Multiple studies argue that the institutional context has an important influence on innovation dynamics and policy effectiveness for knowledge development (Boschma, 2004; Asheim & Coenen, 2005; Rodriquez-Pose, 2013; Boschma et al., 2014; McCann & Ortega-Argilés, 2014; Foray et al., 2018). This is expected to be a two-way relation, where institutions support knowledge development and are changed along the trajectory to keep up with knowledge development. It is therefore important that the institutional context is flexible and can respond to changes in knowledge development (Boschma, 2004). In this study focus primarily lies on the relation of the institutional context in respect to knowledge development given the taken scope on innovation dynamics and policy. With this in mind, the following section explores the possible relations between the two in the regional context.

For knowledge development, a region is dependent on various kinds of mechanisms within and between markets, networks and organisations, which collectively form its institutional context (Boschma, 2004). The functioning of this system is shaped by the manner in which these mechanisms are coordinated and how various parts interact. The tacit nature of, and cultural aspects involved in these processes make it that these mechanisms are strongly geographically bounded (Boschma, 2004; Rodriquez-Pose, 2013). Additionally, these particular elements complement each other in different ways, resulting in the notion that blindly copying institutional element from other regions is to no avail (Hall & Soskice, 2001). Therefore, the institutional context shapes a unique regional environment in which knowledge development takes place. It supports or constrains elements such as trust and transparency, regional variety and mutual learning. Moreover, it impacts the intensity in which the regional interactions manifest and subsequently the degree of learning and knowledge production that takes place (Boschma, 2004). In this regard, previous studies have indicated that a coordinated or liberal political and market structure stimulate incremental or radical knowledge development respectively (Boschma, 2004; Hall & Soskice, 2001). The coordinative and connecting aspects of the institutional context are therefore expected to relate interdependently with knowledge development and the design of territorial strategies (Rodriquez-Pose, 2013).

In the coordinative sense, the government is in a unique overarching position over the knowledge development process (Etzkowitz, & Leydesdorff, 1995). It is in contact with all relevant knowledge development actors and has the tools to tweak the regional knowledge development system. It can do so by imposing its goals on knowledge development organisations or interfere in the science-technology-industry interfaces. Moreover, public bodies are responsible for regulating learning processes, to form mutual trust, facilitate knowledge transmission and provide funding (Rodriquez Pose & Di Cataldo, 2015; Foray et al., 2018). Rodriquez Pose & Di Cataldo (2015) therefore state that the level of knowledge development that a region can achieve is related to the quality of government. They define quality of government as a degree of corruptness, rule of law, accountability and effectiveness of government levels regulating the region.

H2: Quality of governance is positively related to complex knowledge development

The connecting institutional element presents itself in the type of actors present and the manner in which they, individually and collectively, are used to develop knowledge (Boschma, 2004; Asheim & Coenen, 2005; Rodriquez-Pose, 2013; McCann & Ortega-Argilés, 2014; Foray et al., 2018; Trippl et al., 2019). It is a well-established notion that because these factors differ for each region, no one-size-fits-all innovation policy exists. This is the reason why smart specialisation is designed as a bottom-up discovery process for each individual region (Foray et al., 2011). However, previous studies have found that patterns exist between regional interactions and its capability to develop knowledge (Asheim & Coenen, 2005, Boschma et al., 2017). Boschma (2004) states that these regional interlinkages allow knowledge exchange and regional learning to take place. He argues that in general the more regional interlinkages are present, the more flexible the regional system is. Benner (2019) refers to this level of regional interlinkages as the institutional thickness. He argues that the more institutional thickness a region possesses, the more capable it is in understanding its own regional context and in improving it. Asheim & Coenen (2005) state that in this regard first of all, the presence of experimenting actors like universities, research organisations and firms are important in order to perform the learning process. Secondly, the clustering capability of these actors influences to what extent knowledge can culminate and subsequently the level of specialisation and complexity that can be achieved. These findings result in the expectation that institutional thick regions are more likely to produce complex knowledge.

H3: Institutional thickness is positively related to complex knowledge development

Following this discussion, it can be expected that only a few regions possess the optimal regional knowledge-institutional capabilities to diversify into whichever new or complex field of interest. This raises the question how the majority of regions, which do not possess all of these capabilities, strategise in the RIS3 program given their knowledge and institutional context.

2.3. Prioritisation strategy

To find the particular strengths for these regions to focus on in their smart specialisation, the RIS3 prioritisation process was developed (Foray et al., 2011). Foray et al. (2011) describe the prioritisation process as a process of entrepreneurial discovery. In this process, entrepreneurial knowledge of regional actors is consulted, because this includes information about inputs and services such as knowledge on market potential and engineering that other sources do not possess. Note that regional actors include individuals, firms, public organisations and research and education institutions. However, this approach alone leaves no room for preconceived important related topics, while current RIS3 strategies have a more central and active role for the government (Boschma & Gianelle, 2014; McCann & Ortega-Argilés, 2015). Regional governments can fulfil this role by analysing their distinctive scientific portfolio on knowledge characteristics (Heimeriks et al., 2019; Balland et al., 2019). This process of selecting scientific fields based on relatedness and pre-conceived importance by regional governments is labelled as the governmental discovery.

The official RIS3 EU directive explicitly states that entrepreneurial knowledge and a governmental analysis are the key pillars for smart specialisation (European Commission, 2012). Therefore, to include both concepts and remain close to the EU instructions, this study follows Boschma & Gianelle (2014) and McCann & Ortega-Argilés (2015) by conceptualising the prioritisation process as an interplay of governmental and entrepreneurial discovery. Ultimately, the way in which a region tries to achieve their innovation goals is indicated by the collective characteristics of the priority portfolio resulting from the prioritisation process, which is identified as a region's prioritisation strategy (Foray et al., 2018).

Therefore, a region's prioritisation strategy is effectively a selection of knowledge trajectories on which a region wishes to embark. Moreover, given that the prioritisation strategy is thus fundamentally based on the regional context and its capabilities to develop knowledge, this study expects that relations between a region's knowledge and institutional configuration and its prioritisation strategy exist. Although no one-size-fits-all policy would prevail, a similar regional context should therefore result in similar types of strategies (Heimeriks et al., 2019; Balland et al., 2019). Considering the previous discussion, it would be expected that these influencing contextual elements are features of the knowledge base, governmental context and institutional thickness.

2.3.1. Knowledge base

In the pre-described context, Balland et al. (2019) point out that the risk involved in prioritising complexity depends on the relatedness of the topics to the regional knowledge base. Namely, as the RIS3 reasoning suggests, the more related the priorities to the knowledge base, the less risk is involved in the process. Consistent with broader knowledge development theory, this implies that regions that have high diversification opportunities or in other words, a high number of fields that are related to the knowledge base, are more likely to prioritise complex fields due to the limited risk involved. Reasonably, if the options are present, regions are expected to prioritise the fields that have the potential to reap the highest benefits. Similar reasoning applies on the presumed effect of a region's diversification opportunities on the relative relatedness between a region's knowledge base and prioritisation strategy. If all regions look for relatedness in their prioritisation process, then it is expected that regions with more diversification opportunities are relatively better able to connect their priorities to their knowledge base and achieve a higher relatedness to their prioritisation strategy. For regions on the other side of the spectrum, Heimeriks et al. (2019) found that regions that do not possess many diversification opportunities, could be better off by diversifying their knowledge base before attempting to specialise in more complex topics. This would enhance their diversification opportunities in the future and help to reach the level of diversity required to be able to embark on more complex trajectories.

H4a: Related diversification opportunities are positively related to the complexity of the prioritisation strategy
H4b: Related diversification opportunities are positively related to the relatedness of the prioritisation strategy
H4c Related diversification opportunities are negatively related to the diversity of the prioritisation strategy

2.3.2. Governmental context

Like for general knowledge development, the institutional context is also expected to be interdependent with the prioritisation strategy. (McCann & Ortega-Argilés, 2014). Previous studies have pointed out that the smart specialisation program can be beneficial for institutional improvements in streamlining governmental processes and bringing regional actors closer together (Benner, 2019). However, since this study focuses on the prioritisation process, primarily the influences of the regional government and the institutional thickness on the prioritisation strategy are investigated. Governmental quality, from a smart specialisation perspective, is often assessed as the governmental structure in which the prioritisation process takes place (Morgan, 2017; McCann & Ortega-Argilés, 2014). Two main aspects of this structure, reflected in the works of Rodríguez-Pose & Wilkie (2017) and Foray et al. (2018), are expected to influence the effectiveness and direction of the prioritisation strategy.

First, the presence of supportive entrepreneurial and experimental conditions is important as it is a core element of smart specialisation (Foray et al., 2011; Rodríguez-Pose & Wilkie, 2017). Foray et al. (2018) argue that possessing these conditions in a governmental context help governments dealing with their knowledge deficit in novel developments. This deficit is incurred, because the more specialised the innovation becomes, the more knowledge and expertise is required for a competent innovation strategy. To reduce the knowledge deficit, governments will have to show experimental policy behaviour to experiment what public intervention works, how, where and why. Governments could do this for instance by conducting and supporting research and experimenting with policies. It is expected that governments that possess these qualities are better at stimulating variety and therefore prioritise more diversely (Foray et al., 2018; Morgan, 2017). Furthermore, it is expected that with an experimental environment more complex fields can be prioritised, because these qualities reduce the risks involved and improve regional flexibility needed in complex prioritisation (Foray et al., 2018; Balland et al., 2019).

H5a: The public experimental context is positively related to the diversity of the prioritisation strategy
H5b: The public experimental context is positively related to the complexity of the prioritisation strategy

Secondly, efficient interplay between governmental and entrepreneurial actors in the prioritisation process is important (Foray, 2018; Rodriquez-Pose & Wilkie, 2017). However, intimate private-public relations could result in entangled interest and are thus not always beneficial for the prioritisation process (Rodriquez-Pose & Wilkie, 2017). Therefore, an open public engagement climate is needed. Rodriquez-Pose & Wilkie (2017) describe that such a context is mainly established by transparent, non-corrupt governments, that are able to build relationships based on trust. Additionally, Foray et al. (2018) argue that by taking an active role in public procurement, governments can add to this environment. Subsequently, effective public engagement is expected to improve regional actors' understanding of their regional context and therefore increase their capability to steer towards related change (Benner, 2019). Additionally, because developing complexity often involves multiple actors, these more effectively coordinated regions are expected to be more comfortable in prioritising complex fields (Rodríguez-Pose et al., 2014). Moreover, it can be expected that these regions are more capable at choosing the specific complex fields that have a relatively high relatedness, limiting their risks (Balland et al., 2019).

H6a: Effective public engagement has a positive relation to the relatedness of the prioritisation strategy
H6b: Effective public engagement has a positive relation to the complexity of the prioritisation strategy

2.3.3. Institutional thickness
Furthermore, in regard to the institutional thickness, besides governance, at least the higher education sector (HES) and private sector are deemed important for steering the direction and implementation of knowledge development (Etzkowich & Kolfsten, 2005; Carayannis 2014). The HES generates fundamental knowledge, chases unknowns and provides new sources for innovation (Etzkowich & Kolfsten, 2005). The private sector generates applicable knowledge and solves societal needs (Carayannis, 2014). Co-creation among these actors is expected to yield the most fruitful results in terms of knowledge development. However, the representation of these actors within the region could steer the direction of the prioritisation strategy, given the different roles these actors play in knowledge development and their inclusion in the entrepreneurial discovery process (Carayannis, 2014).

With the taken perspective of scientific knowledge as a foundation for regional capabilities, the HES's role should not be overlooked in the prioritisation process. It can be expected that regions prioritise according to HES potential, because the HES often accounts for a substantial part of the regional ability to explore new knowledge trajectories and extend its fundamental knowledge capabilities (Etzkowich & Kolfsten, 2005). Presumably, given the collective nature of the prioritisation process, the representation of the HES in regional knowledge development affects the degree to which these aspects are conveyed to other actors and subsequently the degree to which they are implemented in the prioritisation strategy. Therefore, it is expected that regional HES representation in the prioritisation process steers a region's prioritisation strategy to complex and diverse priorities.

H7a: Regional university representation is positively related to the diversity of the prioritisation strategy
H7b: Regional university representation is positively related to the complexity of the prioritisation strategy

For the representation of the private sector in regional knowledge development, a similar notion could apply. The lobbying activities of the private sector could persuade other actors in the prioritisation process to prioritise in more established fields (Landabaso et al., 2014; Santoalha & Boschma, 2019). It is therefore expected that the more represented industry is in the prioritisation process, the more related the priorities will be.

H7c: Regional private sector representation is positively related to the relatedness of the prioritisation strategy

Lastly, multiple studies point out that in this internal context of regional connectedness and representation, external connections could play an important role in the prioritisation process as well (Goddard, 2013; Rodríguez-Pose & Crescenzi, 2008; Balland & Boschma, 2020). Due to the varying knowledge capabilities among European regions, regions could try to find missing capabilities elsewhere (Balland & Boschma, 2020). This could enable regions to prioritise in fields that their own regional configuration would not allow. Such a scenario is possible in an open region where actors in the prioritisation process possess strong external connections and are aware of capabilities outside of the region. This way regional actors are able to convey more complex strategies in the prioritisation process due to their reliance on extra-regional partners. Additionally, this could imply that these regions are inclined to prioritise more diversely, inspired by extra-regional influences.

H8a: Regional openness is positively related to the diversity of the prioritisation strategy
H8b: Regional openness is positively related to the complexity of the prioritisation strategy

3. Methodology

3.1. Research design and data collection

This study attempts to find answers to the research question in a quantitative cross-sectional research design. As a first step, descriptives were analysed to get a sense of regional distributions. Secondly, the hypotheses on relations between the knowledge and institutional characteristics with knowledge development were assessed in regression analyses. In a third step, expected effects of the regional features deemed relevant for prioritisation on the strategy characteristics were tested in a further set of regression analyses. In this set of regression analyses, the independent and control variables were settled before the dependent variables to ensure that causal relations could be inferred due to the temporal difference (Bryman, 2012). Additionally, regions were clustered and classified based on their level of knowledge-institutional development to assess the institutional attributes present in different types of regions. This was done in order to assess if different knowledge development and prioritisation implications apply to regions in different stages of regional development. In this design, all EU NUTS 2 regions that had available data were included for optimal representability. Additionally, to provide a holistic overview of European knowledge production, all European NUTS 2 regions were included in the descriptive analysis of regional knowledge development.

Data on the scientific knowledge production of the regions was collected with Web of Science publication data. As one of the three major scientific platforms, Web of Science is a credible source for this purpose (Hicks et al., 2015). Although the Web of Science publication data is publicly available via their online services, the Centre for Science and Technology Studies (CWTS) database was used as the data source for this data. The CWTS database provides cleaned bibliometric datasets linked to geographic locations based on spatial coordinates, making it a very accurate data source for this study. For the construction of knowledge base indicators, publication data from 2010 till 2011 was used, preceding the prioritisation period. The corresponding subfield classifications, linked to the publications based on Web of Science journal classification, were used to classify the publications into scientific subfields. The Web of Science divides publications in five main scientific domains that are subdivided in multiple smaller subfields (Web of Science, 2018). A list of all these subfields is provided in Appendix A.

To collect data on the institutional context of European regions, regional databases from Eurostat and the RIS3 Benchmark Regional Structure (BRS) were consulted (Navarro et al., 2014). The Eurostat and BRS databases provide public, private and socio-demographic indicators based on information gathered by the European Commission. Data around the same period as the knowledge base publications ranging from 2010 to 2013, preceding prioritisation, were extracted from these sources. To collect additional relevant data on the regional governance, the articles of Fazekas (2017) and Charron et al. (2014) were used. From these studies data from 2005 till 2015 was collected, remaining within the boundaries of the timeframe applied in this research.

The priorities of the regions were collected via the Eye@RIS3 database (European Commission, 2018a). In this database priority descriptions for European nations and regions reported to the European Commission for the RIS3 program are registered. The years in which the priorities are reported range between 2013 to 2018. These descriptions are registered with corresponding scientific domains, economic domains and policy objectives. In regard to the scientific knowledge perspective taken in this study, the scientific domains were collected alongside the descriptions.

To collect the data in the form of prioritised knowledge trajectories, the priorities were coded along one or more of the Web of Science scientific subfields. This way the correspondence between knowledge capabilities and the prioritisation strategy could be assessed. This was done by a manual process based on an assessment of the priority labels and descriptions provided by the Eye@RIS3 database. In this database a considerable number of smart specialisation priorities were reported on the NUTS 2 level and could thus directly be used for the data collection. However, some priorities were reported on the national, NUTS 1 or NUTS 3 level. In case of a NUTS 3 level registration all priorities registered in a NUTS 3 region, covering the same NUTS 2 region, were collected together as a NUTS 2 level representation. In case of a national or NUTS 1 level registration, the regions were not included in the coding process, because this might have blurred the regional focus, which is at the core of this study.

To ensure validity and replicability, general guidelines were used in the coding process of the priorities. Normally, each registered priority was assigned to two scientific subfields to prevent overrepresentation of one field, when actually the priority fell in between subfields. However, if a priority is aimed towards one specific subfield, only this field was coded to prevent a second field from gaining overrepresentation. Additionally, priorities that were difficult to score and/or were registered multiple times in different regions, were noted with their corresponding codes to ensure that the same procedure was carried out over the whole process. These coding configurations are displayed in Appendix B. Following this procedure, the priority portfolios of 163 regions were collected. The sample was checked on the inclusion of all types of regions in terms of knowledge development to ensure representability. Given that all the data sources discussed above are publicly available, great replicability is provided and therefore the study's reliability is increased.

3.2. Operationalisation

3.2.1. Knowledge development

Complex knowledge development

To operationalise the concept of complex knowledge development, a measure had to be constructed that indicated the complexity of regional scientific knowledge production, based on the obtained publication data. In this context, a two-mode co-occurrence matrix of scientific subfields was constructed, following Balland & Rigby (2017) and Balland et al. (2019), based upon the principles of diversity and ubiquity (Hidalgo & Hausmann, 2009). As Balland et al. (2019) put it figuratively, a Scrabble word with more letters (diversity) and more singular letters (q, x e.g.) (ubiquity), is more complex than a short word with ubiquitous letters. Therefore, this measure showed the level of complexity for each scientific subfield, based on the level of diversity present in the regions in which the field is found and the number of regions a field is found in. Averaging the scores of all the subfields in which a region is producing knowledge, ultimately gave an indication of the complexity of a region's knowledge development

First, a two-mode binary adjacency matrix was constructed (M), based upon in which scientific subfields regions have a revealed comparative advantage (RCA). The RCA was calculated as a binary variable that assumes the value of 1 when a region produced a greater share of publications in subfields i than the average of the EU-28 plus Iceland, Norway and Switzerland as a whole and 0 otherwise. Only fields with an RCA were included to make sure that negligible activity in certain scientific fields does not bias the complexity indicator. Matrix (M) has the dimensions n=282 (NUTS-2) by k=number of scientific subfields.

$$RCA = \frac{Publication_{rit} / \sum_i Publication_{rit}}{\sum_r Publication_{rit} / \sum_r \sum_i Publication_{rit}} > 1$$

To calculate complexity scores of the individual scientific subfields, matrix (M) and its transpose (M^T) were row standardised and multiplied. This resulted in a square matrix (B) with dimensions equal to the number of scientific subfields included.

$$(B) = (M^T * M)$$

Then the scientific complexity index (SCI) for each of the scientific subfields was given by the second eigenvector $\vec{Q}$ of matrix (B) that was standardised as:

$$TCI_i = \frac{\vec{Q} - \langle\vec{Q}\rangle}{stdev\,\langle Q\rangle}$$

The resulting complexity scores for the scientific subfields are listed in Appendix A. Ultimately, to determine the complex knowledge development score for an individual region, the complexity scores of the subfields in which this region has an RCA were averaged. This score will be referred to as the regional complexity. To calculate the values for this indicator, this study made use of the EconGeo package in R (Balland, 2017).

Related diversification opportunities

To construct an indicator for related diversification opportunities, the region-subfields network approach based on Hidalgo et al. (2007) and Heimeriks et al. (2019), was used. As was stated in the theoretical section, there are related diversification opportunities when a new subfield relies on similar knowledge capabilities as subfields that are currently being researched in the region. This branching process can be represented as a network in which subfields are represented as nodes and placed in this network based on the relative relatedness to each other.

In this approach the relatedness between scientific subfield i and j was computed by taking the minimum of the pair-wise conditional probabilities that regions published in one field given that they published in another scientific subfield j $\varphi_{i,j,t}$ during the same period. This implies that scientific fields were assessed as related if they co-occur together in different regions and that the relatedness score ranges between 0 and 1 based on the frequency of these occurrences. To avoid negligible activity only the scientific subfields were included in which regions have an RCA.

$$\varphi_{ijt} = min(P(RCA_{it}|RCA_{jt}), P(RCA_{jt}|RCA_{it}))$$

In order to operationalise related diversification opportunities, the relatedness density was calculated (Boschma et al., 2014). This measure indicates the level of relatedness shown in a region to a specific subfield and therefore gives an indication for the diversification opportunity into this subfield. To calculate the relatedness density score of a scientific subfield i in a region the sum of the relatedness scores from all subfields in the portfolio of the region was taken to subfield i. This was divided by the sum of the relatedness scores of i to the remaining subfields and multiplied by 100 to arrive at a relatedness density percentage. The relatedness density score for the region was then calculated by taking the average of relatedness density scores for all scientific subfields.

3.2.2. Prioritisation strategy

From the theoretical discussion it became apparent that relatedness, complexity and diversity are relevant aspects in scientific knowledge development. Given the knowledge development perspective taken on RIS3 prioritisation, the degree in which regions reflect these principles in their priority portfolio will be an indication of their prioritisation strategy. For this sake, the prioritisation strategy is operationalised as a collection of prioritised scientific subfields.

Complexity of strategy

Based on the complexity scores of the individual scientific fields, a measure for the strategy complexity was calculated. The average complexity score of the prioritised scientific fields for each region were calculated to arrive at the total complexity score for a region's prioritisation strategy.

Relatedness of strategy

The higher the relatedness density of a regional knowledge base to the prioritised subfields, the shorter the figurative distance from its knowledge base to the priority field and therefore the more related the priority is to the region. Therefore, the relatedness density scores calculated for the related diversification opportunities were used to create this relatedness measure. To acquire the relatedness density score for the prioritisation strategy, the sum of the regional relatedness density scores, for the subfields included in its prioritisation portfolio, was taken and averaged. To circumvent the bias that regions with high relatedness density scores will automatically receive high relatedness of strategy scores, these scores were normalised by dividing them with the average regional relatedness density. A benefit of this measure is that it can easily be determined if the prioritisation strategy is more or less related to the regional knowledge base than the average level of opportunity. Namely, if the relatedness of the strategy is above 1, the prioritisation strategy is related above average and if it is below 1, it is related below average. To provide insight in this respect, this in-between measure is displayed in its binary form in Appendix D. However, because there is a significantly higher variance in the relatedness density scores for low relatedness density regions compared to high relatedness density regions, low relatedness density regions score too high or too low with this measure in absolute terms. Therefore, to arrive at the final prioritisation strategy relatedness score, the values were normalised by the inner variance of relatedness density values for each individual region.

Diversity of strategy

To calculate the diversity of the prioritisation strategy the formula from Stirling (2007) was used. This measure was chosen, because it encompasses all three components of diversity, which are variety, disparity and balance, in contrast to other diversity measures. This measure is often referred to as Rao-Stirling diversity in respect to the work of Rao (1982) and Stirling (2007). The formula gives the sum of pairwise disparities, weighted in proportion to contributions of individual system elements.

$$D = \sum_{ij(i \neq j)} d_{ij} * p_i * p_j$$

The factor d_{ij} is the disparity attributed to i and j and the factors p_i and p_j are proportional elements of i and j. Disparity d_{ij} can be calculated as the distance between scientific fields in a network (Leyesdorff & Rafols, 2011). Proportions pi and pj are then assigned based on the number of represented scientific subfields in the prioritisation portfolios. Therefore, the proportion of which a priority portfolio consists of scientific subfield (i,j) is (p). The calculation of these diversity scores for the prioritisation strategy per region was performed based on a region by priority field matrix using the diverse package in R (Guevara et al., 2016). (Guevara et al., 2016).

3.2.3. Institutional context

Quality of governance & institutional thickness

Data on the quality of government from Charron et al. (2014) consists of overall quality of government scores for European nations and regions, comprised by survey data. This data source is often used to indicate the general quality of government and served the same function in this study. To estimate the level of institutional thickness, numbers on public-private co-publication from the BRS database were used, as they give an indication for interorganisational collaboration on knowledge development in the region.

Institutional prioritisation context

The public experimental context was operationalised by obtaining data on the regional R&D expenditure by the public sector from the Eurostat database. This data was used as it gives an indication of the knowledge development efforts put in by the public sector in the region. However, to make it a fair indication of relative public efforts, the data on R&D expenditure by the public sector was normalised by the gross domestic product (GDP).

To measure the effective public engagement, indicators from Fazekas (2017) were used. He measured the quality of governance on public procurement (PP) on a regional level by using data on procurement efficiency, competition, administrative efficiency and government corruption control. These concepts align well with the notions of efficient communication and governmental trust, needed to align with extra-governmental organisations in the prioritisation process and for priority implementation. Therefore, these governance indicators are viable measures for this study. The overall score for public procurement governance is used as an indicator for public engagement throughout the analysis for its overarching representation. The remaining, more specific, indicators are used in the analysis to provide more detail where this is possible without compromise. The indicator of administrative efficiency was not included due to validity concerns raised in the report for this indicator.

The representation of both the private sector and the HES were operationalised by using data on regional R&D investment by the private sector, HES and the public sector from the Eurostat database. Using this data, the relative fraction of R&D expenditure by each of the sectors in the region's total R&D expenditure was calculated. This fraction therefore gives an indication of the share of knowledge development performed by each group of actors (Rodríguez-Pose & Crescenzi, 2008). Given the conceptual set-up of the prioritisation process, it is assumed that the higher the share in knowledge development of a certain sector, the more representation it has in the prioritisation process.

The regional exports as a percentage of GDP from the BRS database was used as an indicator for the regional openness. This is a quantifiable measure that indicates the connectedness and dependence of regional markets with markets outside of the region. Therefore, it gives an indication of the awareness and reliance of regional actors in respect to extra-regional capabilities.

3.2.4. Control variables

In the analysis, control variables were added to control for other regional characteristics that might influence the dependent variables. These are generic regional specifications such as economics (GDP per capita), population, education level and type of governmental structure. They were extracted from the Eurostat database except for the type of governmental structure, indicated by the institutional decentralisation, which was extracted from the BRS database.

3.2.5. Operationalisation table

In this section the operationalisation table that summarises the operationalisation procedure for all variables is presented in Table 1.

Table 1: The operationalisation table which shows all the operationalised variables.

Concept	Dimension	Indicators	Calculation of scores/Source
Complex knowledge development	Complexity of regional knowledge production	Diversity and ubiquity in regional publication data	Combined diversity & ubiquity measure (Continuous)
Complexity of strategy	Relative complexity of the priority portfolio	Average relative level of complexity of the chosen priority fields	Combined diversity & ubiquity measure (Continuous)
Relatedness of strategy	Relatedness of priority portfolio to scientific portfolio	Average distance to the regional knowledge base of the chosen priority fields	Co-occurrence measure (Continuous)
Diversity of strategy	Diversity of the priority portfolio	Level of variety and disparity among the chosen priority fields	Stirling diversity measure (Continuous)
Related diversification opportunities	Relatedness density of the regional knowledge base	Relatedness density of publication data	Co-occurrence measure (Continuous)
Quality of Government	Effectiveness of and trust in the government	Civil perception/action and communication effectiveness	Charron et al. (2014) (Continuous)
Institutional Thickness	Inter-organisational density	Level of organisations and interorganisational relations	BRS database (Continuous)
Public experimental context	Relative knowledge development efforts by the public sector	Governmental R&D divided by regional GDP	Eurostat (Continuous)
Effective public engagement	Quality of public engagement	PP good governance score	Fazekas (2017) (Continuous)
Private sector representation	Relative private sector representation in knowledge development	Private sector R&D divided by total R&D	Eurostat (Continuous)
HES representation	Relative HES representation in knowledge development	HES R&D divided by total R&D	Eurostat (Continuous)
Regional openness	Relative regional openness	Regional GDP divided by regional exports	BRS database (Continuous)
Population	Regional population	Regional population	Eurostat (Continuous)
GDP per capita	Regional economic power	Regional GDP divided by population	Eurostat (Continuous)
Education level	Relative level of education	Percentage of population with tertiary education	Eurostat (Continuous)
Institutional decentralisation	Degree of institutional decentralisation	Degree of institutional decentralisation	BRS database (Continuous)

3.2.6. Correlation Matrix

A correlation matrix was composed in Tables 2 and 3 to show the Pearson correlations among the variables and prevent apparent multicollinearity issues. A notable correlation that could be observed is the high correlation between institutional thickness and GDP per capita of 0,802. This correlation is evidently stronger than the 0,329 correlation of regional complexity with GDP per capita, between which a high correlation would be expected. This might hint that institutional thickness plays a notable role in generating regional wealth or vice versa. Overall, no noticeably high correlations could be observed between independent variables included in the same regression model. More details on the means, standard deviations and histogram distributions for the variables can be found in Appendix C.

Table 2: First part of the correlation matrix

Variable	#	1	2	3	4	5	6	7	8	9	10
Regional complexity	**1**	1	0,160	0,825	0,082	0,395	0,668	0,260	-0,131	0,469	0,194
Strategy complexity	**2**	0,160	1	-0,117	-0,239	0,307	0,216	0,188	-0,009	0,104	-0,080
Strategy relatedness	**3**	0,825	-0,117	1	0,177	0,496	0,579	0,062	-0,052	0,586	0,128
Strategy diversity	**4**	0,082	-0,239	0,177	1	0,082	-0,015	-0,099	-0,157	0,159	0,094
Relatedness density	**5**	0,395	0,307	0,496	0,082	1	0,558	0,392	0,015	0,610	0,209
Quality of Government	**6**	0,668	0,216	0,579	-0,015	0,558	1	0,458	0,123	0,589	0,395
Institutional Thickness	**7**	0,260	0,188	0,062	-0,099	0,392	0,458	1	0,183	0,144	0,385
Public experimental context	**8**	-0,131	-0,009	-0,052	-0,157	0,015	0,123	0,183	1	-0,071	-0,122
Effective public engagement	**9**	0,469	0,104	0,586	0,159	0,610	0,589	0,144	-0,071	1	0,132
Private sector representation	**10**	0,194	-0,080	0,128	0,094	0,209	0,395	0,385	-0,122	0,132	1
HES representation	**11**	0,014	0,158	-0,010	-0,151	0,052	-0,033	-0,263	-0,186	0,044	-0,680
Regional openness	**12**	0,015	-0,172	-0,105	0,053	-0,224	0,127	0,153	-0,114	-0,200	0,161
Population	**13**	-0,129	-0,012	-0,173	0,281	0,085	-0,325	0,108	-0,337	0,015	0,196
GDP per capita	**14**	0,329	0,181	0,144	-0,217	0,313	0,550	0,802	0,456	0,165	0,418
Education level	**15**	0,452	0,254	0,471	-0,091	0,664	0,623	0,548	0,068	0,583	0,283
Institutional decentralisation	**16**	0,036	-0,055	0,141	0,002	0,147	0,016	0,008	0,194	0,370	0,028
PP competition	**17**	0,334	0,134	0,485	0,102	0,516	0,345	0,096	0,029	0,805	0,092
PP control of corruption risks	**18**	0,445	0,249	0,431	0,133	0,669	0,622	0,359	-0,039	0,867	0,183
PP administrative efficiency	**19**	0,461	-0,040	0,634	0,177	0,448	0,474	-0,016	-0,038	0,825	0,115

Table 3: Second part of the correlation matrix

Variable	#	11	12	13	14	15	16	17	18	19
Regional complexity	**1**	0,014	0,015	-0,129	0,329	0,452	0,036	0,334	0,445	0,461
Strategy complexity	**2**	0,158	-0,172	-0,012	0,181	0,254	-0,055	0,134	0,249	-0,040
Strategy relatedness	**3**	-0,010	-0,105	-0,173	0,144	0,471	0,141	0,485	0,431	0,634
Strategy diversity	**4**	-0,151	0,053	0,281	-0,217	-0,091	0,002	0,102	0,133	0,177
Relatedness density	**5**	0,052	-0,224	0,085	0,313	0,664	0,147	0,516	0,669	0,448
Quality of Government	**6**	-0,033	0,127	-0,325	0,550	0,623	0,016	0,345	0,622	0,474
Institutional Thickness	**7**	-0,263	0,153	0,108	0,802	0,548	0,008	0,096	0,359	-0,016
Public experimental context	**8**	-0,186	-0,114	-0,337	0,456	0,068	0,194	0,029	-0,039	-0,038
Effective public engagement	**9**	0,044	-0,200	0,015	0,165	0,583	0,370	0,805	0,867	0,825
Private sector representation	**10**	-0,680	0,161	0,196	0,418	0,283	0,028	0,092	0,183	0,115
HES representation	**11**	1	0,073	-0,233	-0,361	-0,204	-0,188	-0,105	0,038	-0,060
Regional openness	**12**	0,073	1	-0,109	0,090	-0,100	-0,340	-0,167	-0,227	-0,228
Population	**13**	-0,233	-0,109	1	-0,030	-0,010	0,199	0,219	-0,008	0,041
GDP per capita	**14**	-0,361	0,090	-0,030	1	0,510	0,176	0,174	0,326	0,108
Education level	**15**	-0,204	-0,100	-0,010	0,510	1	0,302	0,587	0,579	0,494
Institutional decentralisation	**16**	-0,188	-0,340	0,199	0,176	0,302	1	0,432	0,187	0,605
PP competition	**17**	-0,105	-0,167	0,219	0,174	0,587	0,432	1	0,549	0,716
PP control of corruption risks	**18**	0,038	-0,227	-0,008	0,326	0,579	0,187	0,549	1	0,560
PP administrative efficiency	**19**	-0,060	-0,228	0,041	0,108	0,494	0,605	0,716	0,560	1

3.3. Imputation

As can be seen in Figure 1 data was missing for some regions on institutional indicators. Multiple imputation was used in SPSS to create complete datasets that were useful for some analyses. Namely, the construction of networks and clusters required complete datasets and it provided an extra robustness check for the regression models. Given that sufficient data was available for the analyses carried out for all European regions, multiple imputation was only performed for the 163 regions of which the prioritisation strategy was collected. Multiple imputation with chained equations was selected as a valid method for this purpose, since the data had an arbitrary distribution of missing values (Rubin, 1996). However, due to the reasonable amount of missing data, the imputed regression models were not regarded as the main models, considering the inaccuracies that could be involved. First, plausible values were generated using an iterative process of the Markov chain Monte Carlo (MCMC) of five iterations (IBM, 2020). In each iteration a missing value was predicted by all other values in the variable using a univariate model. Due to the random element in the process this created five complete datasets with slightly varying values for the missing data. These datasets were pooled into one complete dataset using the OMS panel, choosing tables, frequencies and then statistics (van Ginkel, 2014). Additional statistics from this procedure are shown in Appendix E

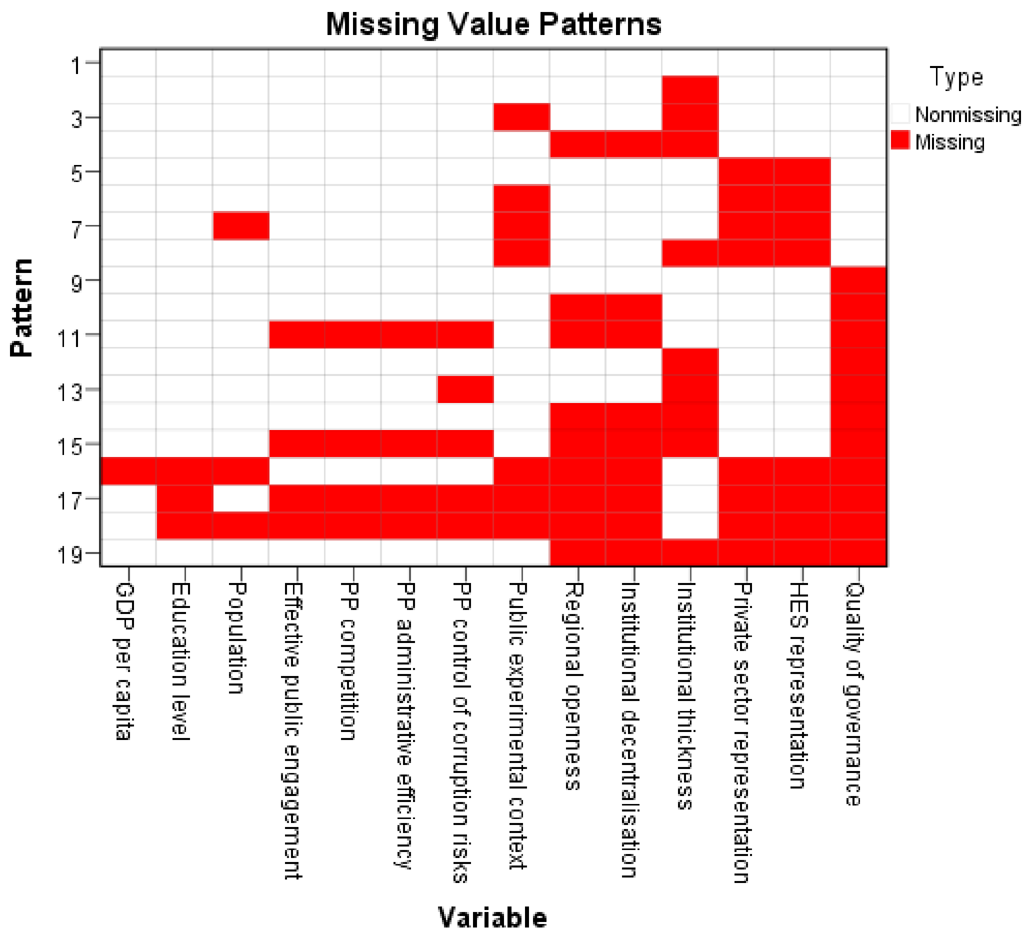


Figure 1: Overview of missing data patterns for the regions subjected to the imputation procedure

3.4. Data analysis

As a first step, insightful descriptives were generated and analysed to be able to place the following analyses in the right context. This was done by qualitatively analysing the descriptives based on the theoretical discussion, in order to infer some preliminary understanding. To test for relations in both steps of the analysis, multiple regression (MR) was performed using SPSS. MR was the best fit for this purpose since the dependent variables all have continuous values (Rubinfeld, 2000). Before constructing the model, the data was checked on outliers and normal distribution to prevent violating the regression principles (Rubinfeld, 2000). Lastly a cluster analysis was performed and qualitatively assessed.

For the descriptive analysis, map and network visualisations were constructed using the Eurostat and Ggmap R-packages and VOSviewer respectively (van Eck & Waltman, 2013). In the map visualisations, the indicators constructed were shown across European regions and qualitatively discussed. Networks were produced in the form of a scientific, institutional and priority relatedness space. The scientific relatedness space was constructed based on the co-occurrence matrix produced by calculating the regional relatedness density scores. In this space, scientific subfields that were often found in the same region were placed closely together Therefore, it provides insight in the degree of relatedness between scientific subfields in a European context.

Additionally, for the 163 regions of which the prioritisation strategy was collected, a measure of institutional density was constructed with imputed data to give an overview of the distribution of institutional features across European regions. In the construction of this measure, all the institutional elements and features in the operationalisation table plus the specific PP governance indicators were included. The institutional density was calculated similarly to the binary methodology of the relatedness density. When regions score an above average value for an institutional attribute this was scored as a 1 and below average as a 0. Then the institutional density was calculated following the same co-occurrence methodology as was used for the relatedness density. With this measure, a distribution of institutional density across European regions and an institutional relatedness space could be constructed. The institutional relatedness space was constructed in the same way as the scientific relatedness with the same analytical purpose in mind. Additionally, in this space, the regional elements GDP per capita and education level were included to provide insight into how economics and education relate to the institutional features.

The priority relatedness space was constructed by assessing prioritised subfields with a 1 and the remaining subfields with a 0. The resulting data frame was converted in a co-occurrence matrix to construct the priority relatedness space, as it was done for the other relatedness spaces. This space provides insight in which priorities are often found together in the same region and therefore in the relatedness among priorities. To cluster the datapoints presented in the relatedness spaces, based on co-occurrence, the unified clustering method of VOS mapping was used (Waltman et al., 2010). This aids in revealing which elements generally group together.

In the regression analysis, four models were constructed in total, one for the dependent variable regional complexity in step one and three for each dependent strategy characteristic in step two. The coefficients calculated in the models were considered statistically significant when their p-value is lower than 0,05. The first model was run without the inclusion of an imputed data model, since enough data was available. This model was also run without the inclusion of the control variable institutional decentralisation, because this lost a lot of cases due to missing data. When no relation was found for this variable and no significant model deviations were observed after exclusion, it was omitted from the model. The imputed models, in the second phase of regression models, were labelled as “full model”, where the model constructed by original data was labelled “control model”. If large variations between the two models were observed, the control model was considered leading and the variations assessed accordingly.

The quality of these models was assessed by calculating the statistical power in the form of the R^2 and adjusted R^2 -value and interpreting these scores. A value below 0,1 is considered as problematically low and was assessed accordingly during the analysis. An additional multicollinearity check was performed by calculating the VIF-scores. If a variable showed a VIF-score above two, the case was investigated to see if problematic multicollinearity issues were present. Presence of problematic outliers was assessed by plotting the residuals, which were judged according to their homoscedasticity. By adding these quality indicators, the study ensures that it upholds its reliability and validity. The quality indicators for the regression models are listed in Appendix F.

In addition to the regression models, clustering analyses were performed to get insight in the distribution and relations of several of the study’s aspects in respect to the level of regional development. There are multiple methods for clustering data available (Fraley & Raftery (1998). One that is often used to form data point clusters is K-mean clustering (Likas et al., 2003). This type of clustering was used when specific data points needed to be clustered, which was the case for classifying the regional types. This analysis aimed to provide insight in the distribution of the discussed regional attributes in different types of regions based on their knowledge-institutional configuration. Therefore, it would provide insight on how different types of regions are organised and what kind of effect this has on knowledge development and prioritisation. Three clusters were formed for this purpose, because creating a lagging, intermediate and advanced cluster made the most analytical sense looking at previous theory (McCann & Ortega-Argilés, 2015; Heimeriks et al., 2019). Moreover, this division was the most statistically sound with quite an even spread over the clusters and significant differences between clusters as is shown in Appendix G. The regional types were determined by the regional complexity, institutional thickness and quality of government as they largely set the knowledge-institutional configuration for knowledge development, following the theoretical discussion. This regional classification enabled a second K-mean clustering with all institutional features included. This resulted in a plot of the distribution of all institutional features in the three regional types. Before analysing the second cluster analysis, it was checked if regional types were reasonably similar to ensure comparability between the two cluster analyses. To perform the K-mean clustering, the built in SPSS function for K-mean clustering was used. This algorithm calculates the Euclidian distance based on the sum of squares and forms clusters based on these distances between data points.

4. Results

4.1. Descriptive analysis

4.1.1. Regional knowledge development

The number of publications on a regional level in the years 2010 and 2011 are displayed in Figure 2 on the left to get a sense of the distribution of knowledge production in Europe. In general, large differences and a high skewness in the values of the highest class are visible in neighbouring regions compared to the lower classes. The skewness is prevalent given that the most publishing region in the second class only reaches 11.729 publications, where the top region in the high class (Ile-de-France) scores ten times higher. The dominance of metropolitan areas in this regard is shown by Figure 3, which displays the top ten publishing regions. These are all metropolitan areas, confirming the importance of proximity advantages for the circulation of ideas and creation of novelty (Balland & Rigby, 2017; Heimeriks et al, 2019). This metropolitan trend remains visible even when the number of publications is controlled for by population size.

However, these numbers do not yet completely assess the quality of the knowledge production in the regions. The map on relatedness density on the right of Figure 2 and the map on knowledge complexity in Figure 5 provide more insight in this regard. The relatedness density gives an indication of the average level of relatedness that the regional knowledge base possesses to any given scientific subfield. Therefore, it provides insight in the ability of regions to connect to new fields and the number of options for low risk diversification and prioritisation a region has.

Looking at the right map in Figure 2, the dispersion of relatedness density across European regions is not as striking as for the number of publications, but it is still visible. The higher scores are somewhat more concentrated in north-western Europe and Spain, compared to the number of publications. The clear differences in regional scores between the two also indicate that the opportunities for related diversification are not just a matter of the quantity of scientific output. This further implicates that the quantity of scientific output does not fully dictate a region's smart specialisation potential (Boschma & Gianelle, 2014). To explain why regions show better or worse potential than would be expected from knowledge output, previous studies have pointed at a region's institutional context (Boschma et al., 2014; Balland et al, 2019; Heimeriks et al, 2019). However, they have not focused on determining institutional specifics in this regard, what this study aims to do in the proceeding analyses.

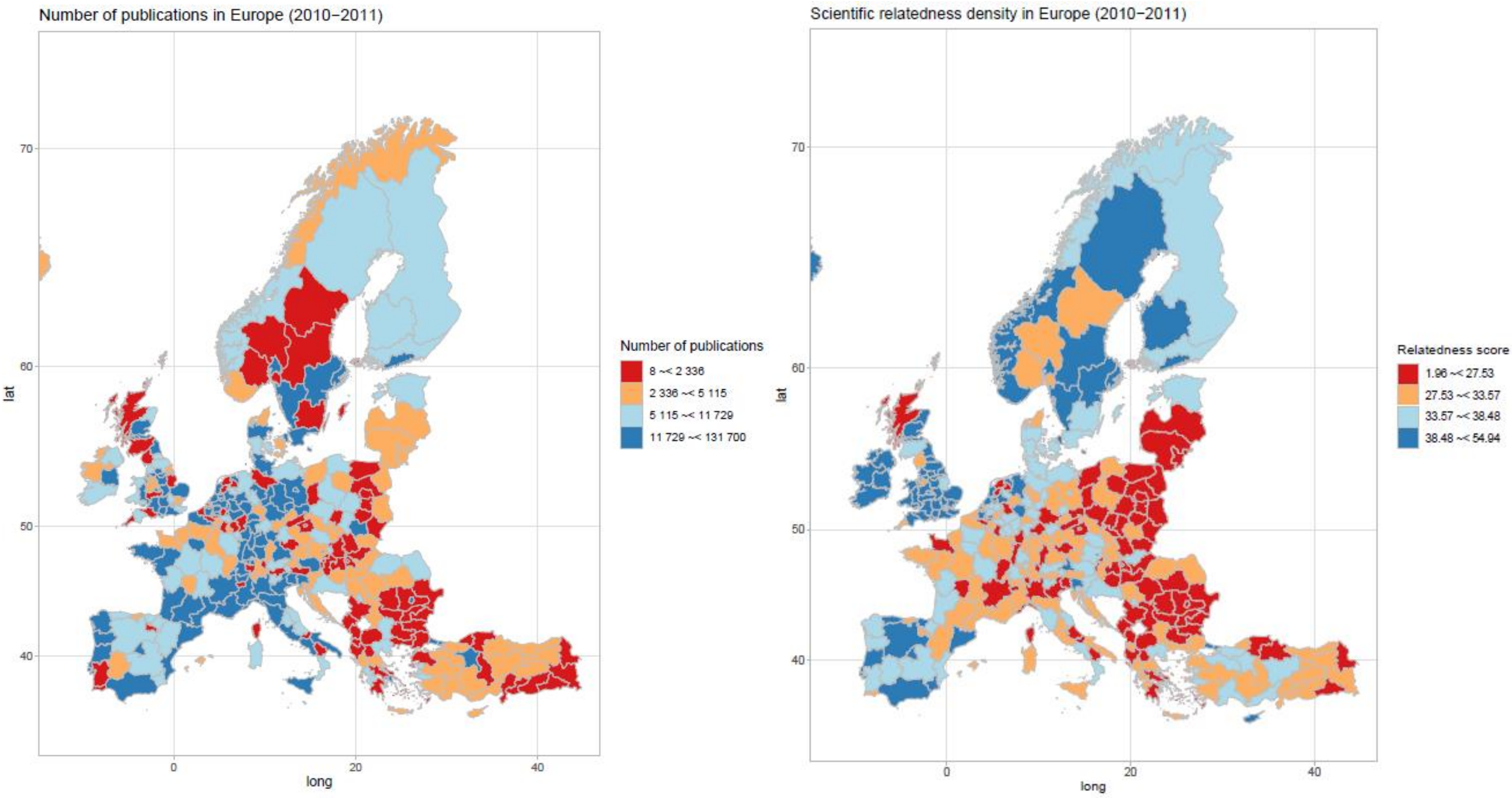


Figure 2: Maps showing the total number of publications across Europe (2010-2011) (on the left) & the relatedness density across Europe (2010-2011) (on the right).

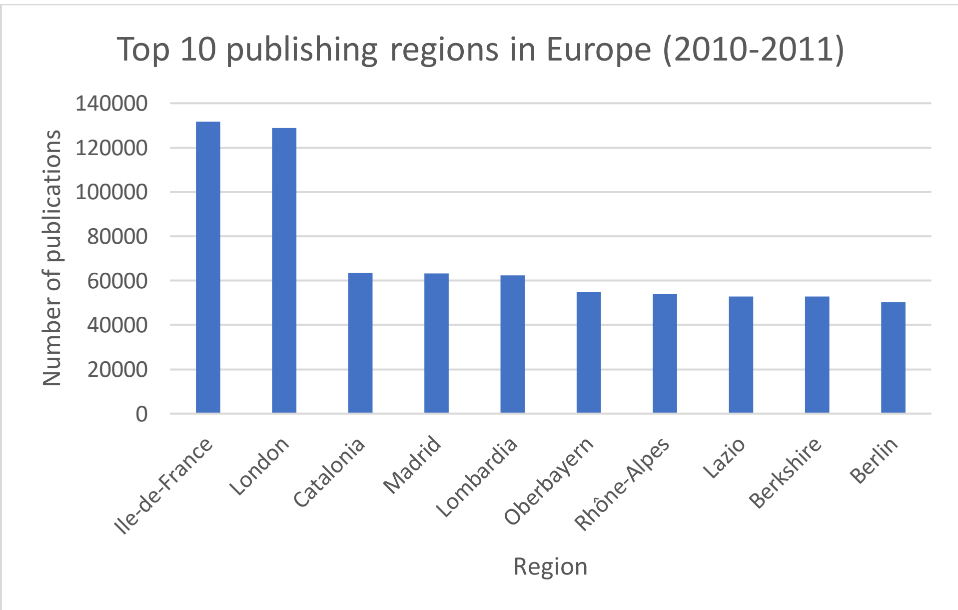


Figure 3: Top ten publishing regions in Europe.

The complexity score gives a representation of how knowledge demanding and unique it is to be active in a certain scientific subfield and therefore indicates the level of knowledge development taking place in the region. The top 15 scientific subfields in this regard are shown in Figure 4 and a list for all scientific subfields can be found in Appendix A. Predominantly social subfields can be found in

Figure 4, as these are likely of multidisciplinary nature and not abundantly present across European regions. In the lower spectrum of complexity fields in physics, chemistry, manufacturing and agriculture can be found, relative basic fields in which a lot of regions are operating.

A European distribution of the regional complexity, taken as the average complexity score of the scientific subfields in which a region is active, is shown in Figure 5. Here, a clear divide across Europe is visible in contrast to the maps shown in Figure 2. Mainly regions in north-western Europe and Switzerland score high compared to other European regions. The complexity measure deviates even more than the relatedness density did to the number of publications, supporting the expectation that reaching complexity does not solely rely on knowledge production. Figure 5 shows a slight preference towards metropolitan regions, but much less than is found in Figure 2 . Especially interesting are some notable differences between regional relatedness density and complexity. Spanish regions show high relatedness density, but fall behind in terms of complexity. This could be an indication that Spain has high complexity development potential, but has not been able to utilise this yet or could be developing this in coming years. On the contrary, Swiss regions show relatively low relatedness density, but high complexity. This shows that, although it is rare, specialised knowledge bases that do not rely on broad knowledge capabilities do exist. However, these regions will likely find it difficult to adapt to change when circumstances call for it, as they lack the low risk potential to connect easily to other fields (Balland et al., 2019).

These results suggest that the level of knowledge complexity is even less steered by the mere knowledge generation capacity and is not completely compliant with related diversification opportunities either. Some regions with low knowledge capabilities show the ability to sustain complex trajectories, where regions in other parts of Europe with higher knowledge production cannot. This is another indication that other (institutional) determinants are involved in these knowledge development dynamics.

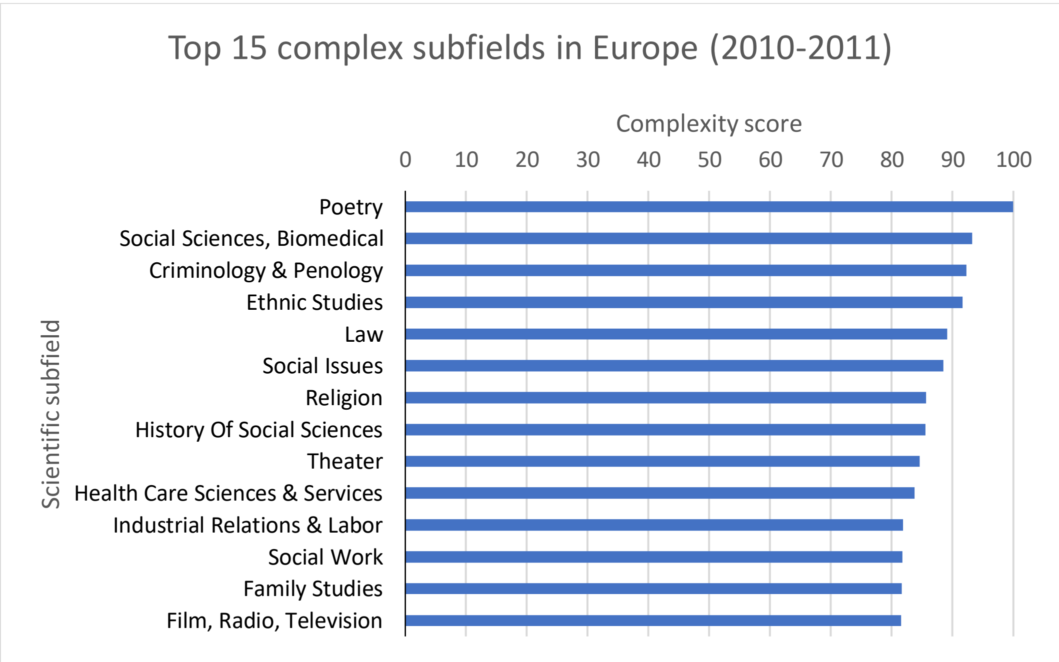


Figure 4: Top 15 complex scientific subfields in Europe (2010-2011)

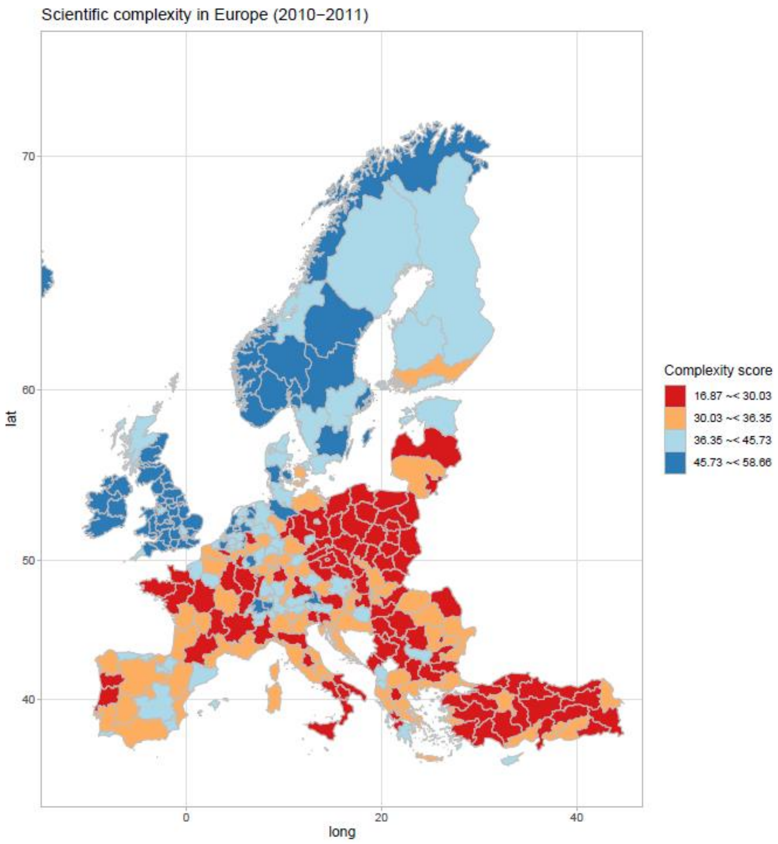


Figure 5: Map showing the regional scientific complexity across Europe (2010-2011).

To provide context to the content of regional knowledge production and its relations, Figure 6 was constructed. Here a scientific relatedness space is presented based on the regional publication data in the corresponding subfields. The more often an RCA of scientific fields is found together in the same European region, the closer these scientific fields are presented in this space. If a line is drawn between two fields, this represents a relative high relatedness between these fields. Additionaly, clusters are formed based on the relatedness scores among the scientific fields. These clusters show a clear division in an engineering focused cluster in green, a social sciences focused cluster in red and a health focused cluster in blue. This indicates that finding high capabilities for these three broad regions in the same region is highly unlikely and that regions typically specialise in one of the three. For the prioritisation strategy, it implicates that prioritising in all the three clusters will result in the most potential to acquire regional diversity, but acquiring high capabilities in all three is inconceivable for most regions. Most regions therefore have to consider if they want to pursue mainly specialisation or diversification in their prioritisation strategy (Heimeriks et al., 2019).

Especially interesting is the high number of clustering lines in the engineering cluster and the lack thereof in the health and social sciences sector. This indicates that high capabilites in engineering fields are often found together in the same region, where specialisation in multiple medical and social fields is more spread out over different regions. An explanation for this distribution can be found in the complexity scores where social sciences score relatively high and engineering sciences relatively low. This shows that engineering capabilities are more often found in European regions in general and therefore more likely to be in the same region as well. Additionaly, the complexity score indicates more interdisciplinary requirements for social sciences, which are not as likely to be found in the same region. Besides being relatively complex and multidisciplanary, the health sector has a relative high resource dependency that could further explain why a large number of health subfields cannot be sustained in a single region. In terms of knowledge development this would imply that, diversifying into other engineering fields, when an engineering basis is present, would be more likely to succeed compared to the other two broad fields, since additional pre-existing conditions are required. From an institutional perspective, this implies that for developing engineering knowledge intra-industry institutions are relatively more important, where inter-industry institutions are relatively more important for the development of health and social sciences. Eitherway, it shows an important role for the institutional context in how to furnish actor interactions and knowledge exchange to enable the development of knowledge trajectories in different industries.

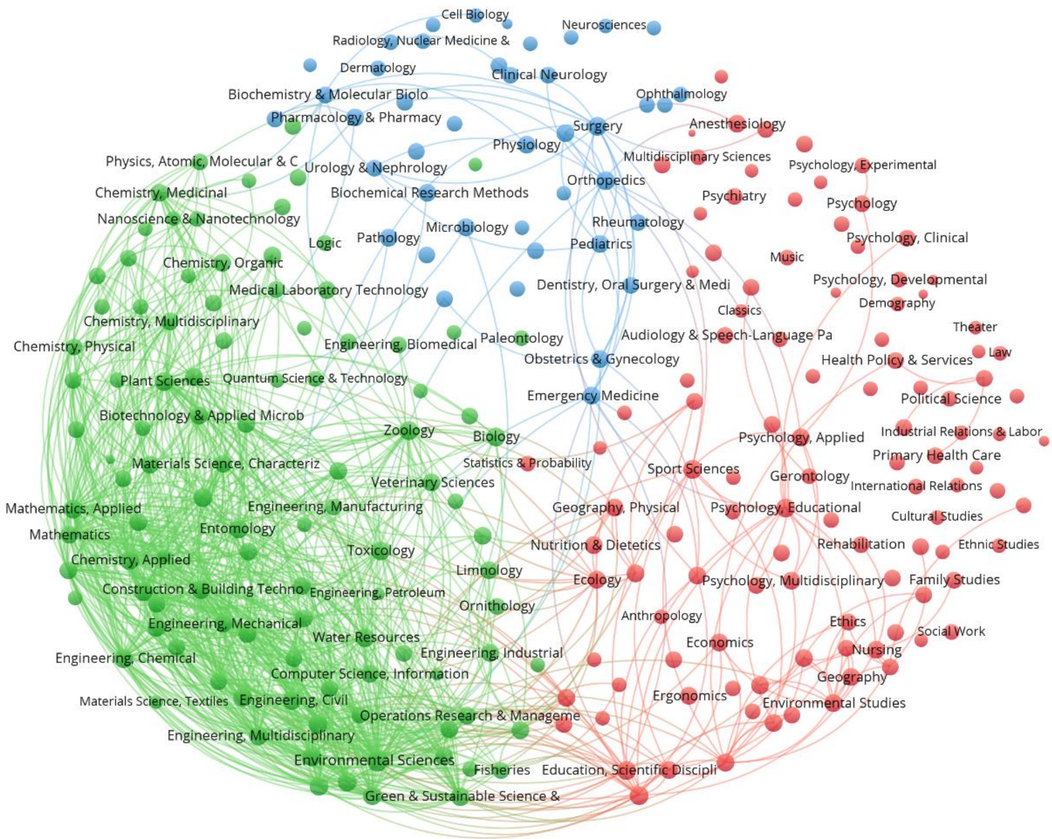


Figure 6: Scientific relatedness space showing the relatedness across scientific subfields based on co-occurrence in European regions (2010-2011).

4.1.2. Regional institutional context

To provide insight on the European distribution of institutional features deemed relevant for knowledge development and prioritisation, Figure 7 was constructed. It shows a similar measure for institutional capabilities as the relatedness density did in Figure 2 for the knowledge base. Therefore, the institutional density gives a representation of the level of institutional capabilities present in the region. Relatively high levels of institutional density can be seen in north-western Europe, France and Austria, where lower levels of institutional density can be found in south and eastern Europe. In this regard it shows quite some resemblance with the distribution of knowledge base complexity in Europe. Most interesting is that some regions, like the French and Spanish regions, meet on about level footing in terms of knowledge base complexity. However, the Spanish regions show higher knowledge base relatedness and the French regions show higher institutional density than the other. This could mean that the French regions are able to compensate for their lower knowledge base relatedness density due to their higher level of institutional capabilities. Still, these are promising showings for the inclusion of these institutional features in knowledge development dynamics.

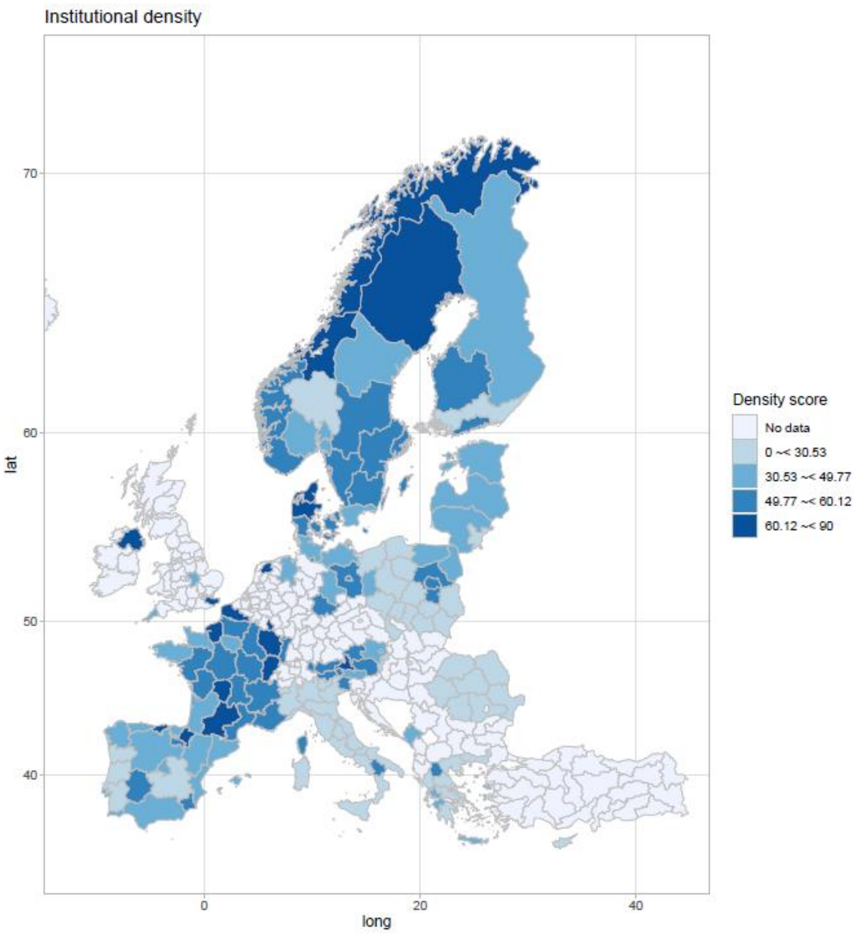


Figure 7: Map showing the institutional density in European regions (2010-2013)

Based on the same principles as used in Figure 6, an institutional relatedness space was constructed shown in Figure 8. This provides an overview of which institutional features typically coexist in regions and which do not. Additionally, the regional elements GDP per capita and education level were included to provide insight into how economics and education relate to the institutional features.

The red cluster seems to be represented by institutional features associated with the structure in which knowledge development takes place, of which the patterns shown are not very surprising. The PP governance indicators are grouped together and institutional thickness is placed among associated institutional features of private sector representation, quality of government, GDP per capita as well as the governmental experimental context. Lastly, the quality of government takes a central place within the structural features as would be expected from its coordinative role (McCann & Ortega-Argilés, 2014).

The green cluster seems to be formed by the institutional features indicating creativity and openness. Namely, institutional decentralisation and open export indicate an open regional structure, and public experimental context and HES representation indicate a source of experimentation. On an individual level, it is interesting that HES representation does not seem to be highly related to education level and GDP per capita, implicating that in less developed regions the HES has a relatively large responsibility in knowledge development. The private sector seems to take over some of this responsibility in the more developed regions. Also very interesting is the surprising distance between public experimental context to quality of government and its high relatedness to private sector features. The ability or willingness of regional governance to experiment therefore seems to be more associated with the presence of a creative and open private sector than the overall quality of government.

The distinction observed between the open creative and structured cluster show resemblance with distinctions in liberal versus coordinated markets and political ideologies. Therefore, if relations between these features and the prioritisation strategy were to be found, it could implicate that a region's liberal or coordinated orientation influences its institutional capabilities. Subsequently, it could provide an explanation to why the types of knowledge development (incremental versus radical) differ between differently coordinated regions (Boschma, 2004).

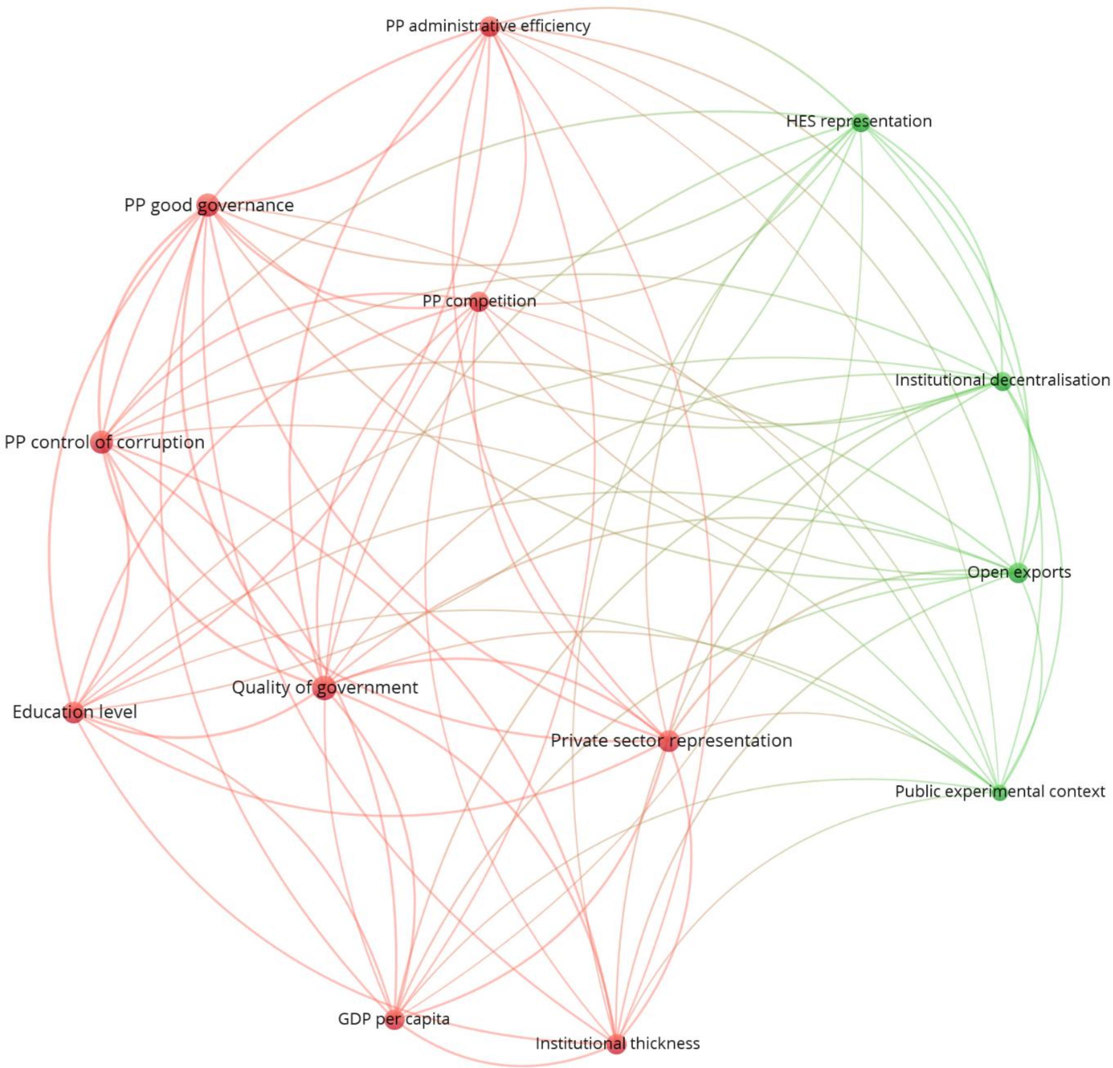


Figure 8: Institutional relatedness space showing the relatedness across institutional features based on co-occurrence in European regions (2010-2011).

4.1.3. Prioritisation Strategy

This section provides more insight on the type of knowledge development pursued by the prioritisation strategies of the European regions participating in the RIS3 program. Three maps are displayed in Figure 9, 10 and 11 to show the distribution of the prioritisation strategy measures across Europe. Figure 9 shows the complexity of the prioritisation strategy based on the constructed priority scores. Resemblance between the complexity of the knowledge base and that of the prioritisation strategy is partly shown by the northern countries that score high and some eastern European regions that score low in both regards. Furthermore, in most countries the metropolitan regions prioritise more complexity although there are certainly exceptions. Overall, large variations are visible within countries and between neigbouring regions.

It is remarkable that there are quite some differences between the complexity of the regional knowledge base and complexity of the prioritisation strategy. This means that some regions prioritise above or below their current level of complexity, which could either prove risky or a waste of potential . An explenation for relative low complexity prioritisation could be that these regions are looking for new pathways or old capabilities that were lost in their trajectories and are currently societaly desirable. Besides opportunistic behaviour, relative high complexity prioritisation could be explained by favourable underlying capabilities like knowledge relatedness density and institutional features.

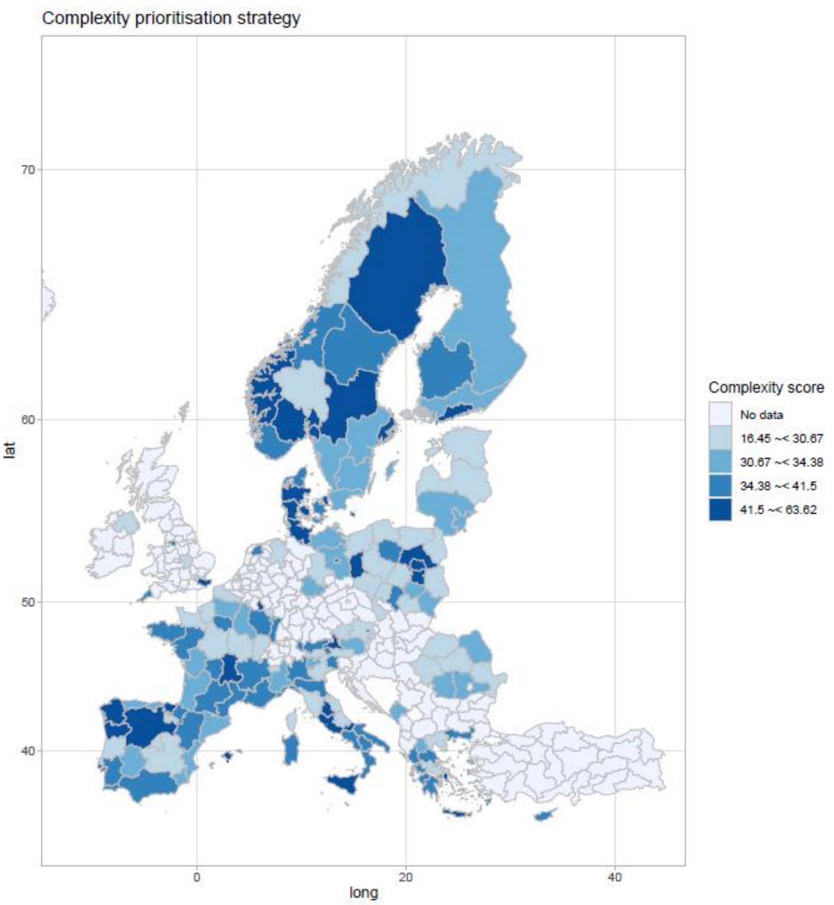


Figure 9: Map showing the complexity of the prioritisation strategy in European regions in the RIS3 program.

Figure 10 shows the relative relatedness of the prioritisation strategy in relation to the regional knowledge base. For this prioritisation attribute, neighbouring regions score quite similar. The northern regions show high values overall, just as they did for the prioritisation complexity. This indicates high knowledge capabilities in these regions, given that they can prioritise both complex and related, so aiming for high gains with limited risks. Besides these regions, the Spanish regions also show high strategy relatedness, presumably relying on their high relatedness density. Regions with lower prioritisation relatedness can be primarily found in Poland, France and Italy, implicating that these regions take more risk in their prioritisation strategy. In general, quite some resemblance can be seen with the knowledge base relatedness density, indicating that regions with more related options to choose from have an improved ability to generate relatedness in their strategy.

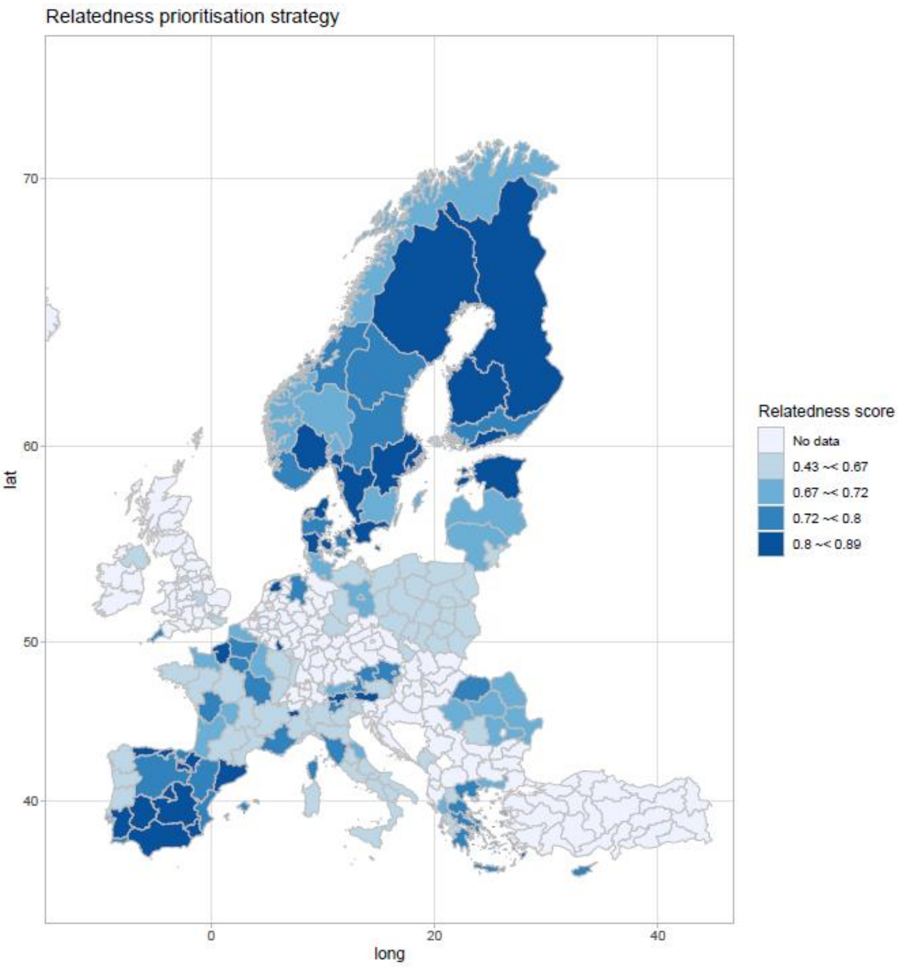


Figure 10: Map showing the relative relatedness of the prioritisation strategy in European regions in the RIS3 program.

Figure 11 shows the diversity of the prioritisation strategy based on the constructed Rao-Stirling diversity scores. Similar to the complexity of the prioritisation strategy, this map also shows large differences within countries and between neighbouring regions. The seemingly random distribution could indicate that the diversity of the prioritisation strategy is highly dependent on the manner in which the prioritisation process is organised on an individual level. It could also indicate that there is not much attention for strategy diversity within the prioritisation process.

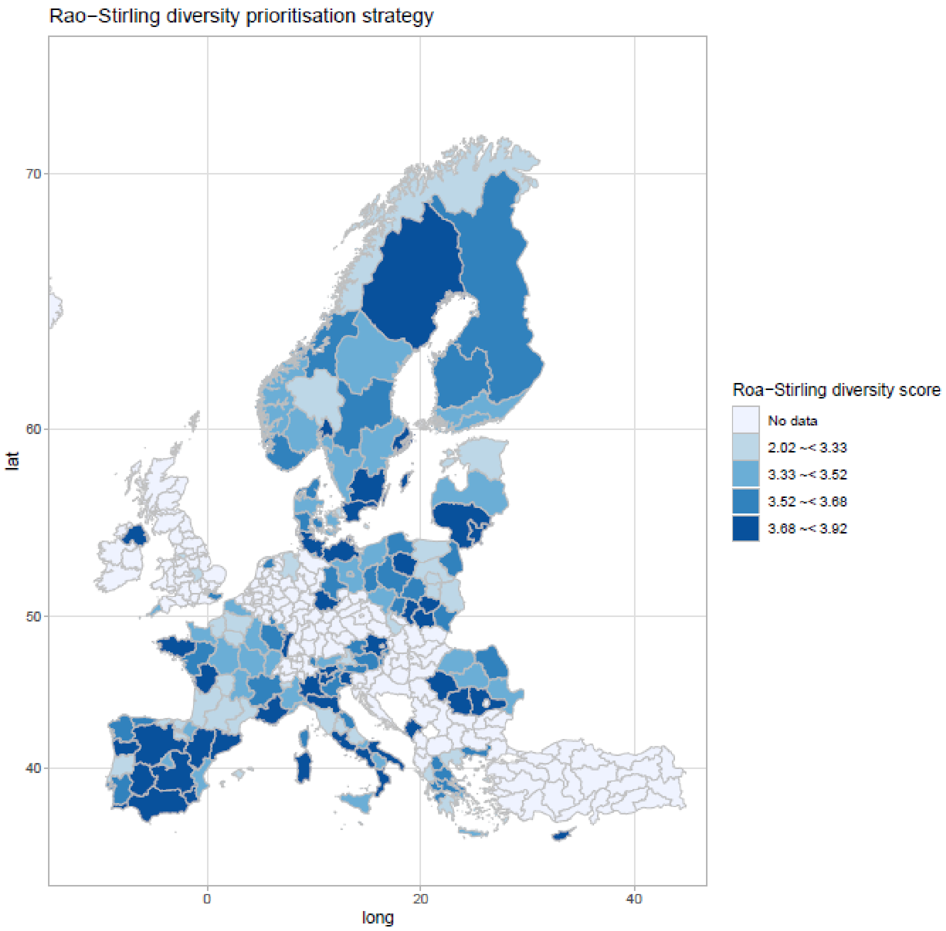


Figure 11: Map showing the diversity of the prioritisation strategy in European regions in the RIS3 program.

The following figures illustrate which subfields were targeted by the European regions and how they relate in a broader context. Figure 12 provides an overview of the number of times a specific scientific subfield is included in a region's priority portfolio. Note that due to the large number of subfields, not all field labels could be presented in one graph, but for the purpose of showing the distribution it is sufficient. It mainly illustrates the highly skewed distribution of scientific subfields, since half of the scientific subfields were not selected ones, compared to the top subfields being selected 100 times. This shows that there is a selective bias among the regional actors for a certain group of scientific fields. Given the smart specialisation rationale and its goals, this is probably due to a combination of a scientific field having a high societal relevance and a relative high level of relatedness to a high number of regions.

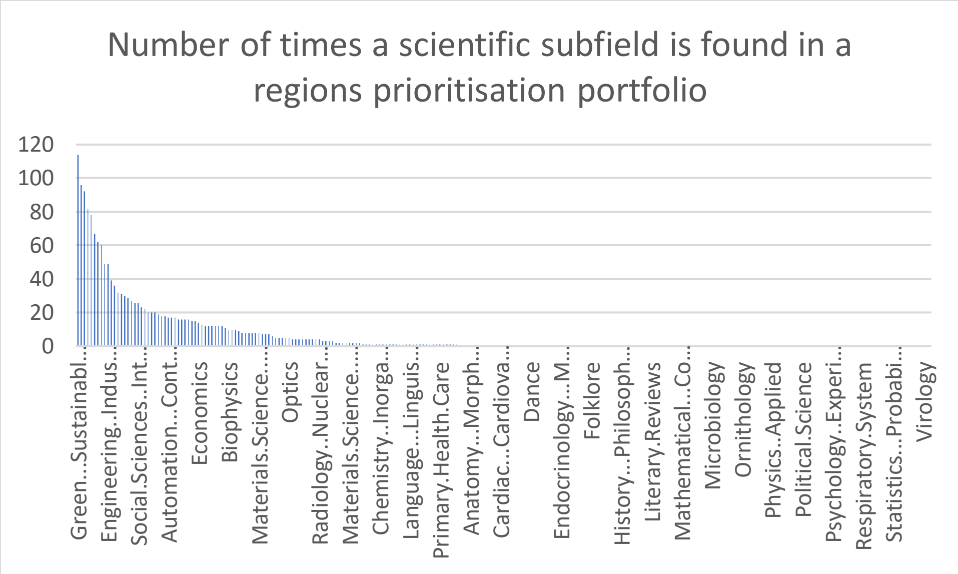


Figure 12: Chart showing the distribution of the times a scientific subfield is selected as a priority in the RIS3 program.

Figure 13 gives a clearer indication of which fields were popular by showing the top ten selected priority fields. Striking is that the socially relevant topics of sustainability and health rank highest, as well as the high position for upcoming sciences and technologies like computer science. While this makes it understandable that regions prioritise these topics, it could be questioned if all regions possess the necessary capabilities, from a knowledge and institutional perspective, to diversify into these fields. Especially for health-related fields this could be problematic, given the high level of complexity associated with these fields. The popularity of food science and engineering fields could be expected, because a lot of regions show to practice agriculture and manufacturing, which make these fields abundantly related across Europe. Overall, mostly general fields are found in the top ten, which is not surprising either, because these are probably widely applicable to many regions compared to the more specific fields. However, it could be questioned how effectively regions can follow-up these priorities in the implementation phase when not much specificity is provided (Foray et al., 2018).

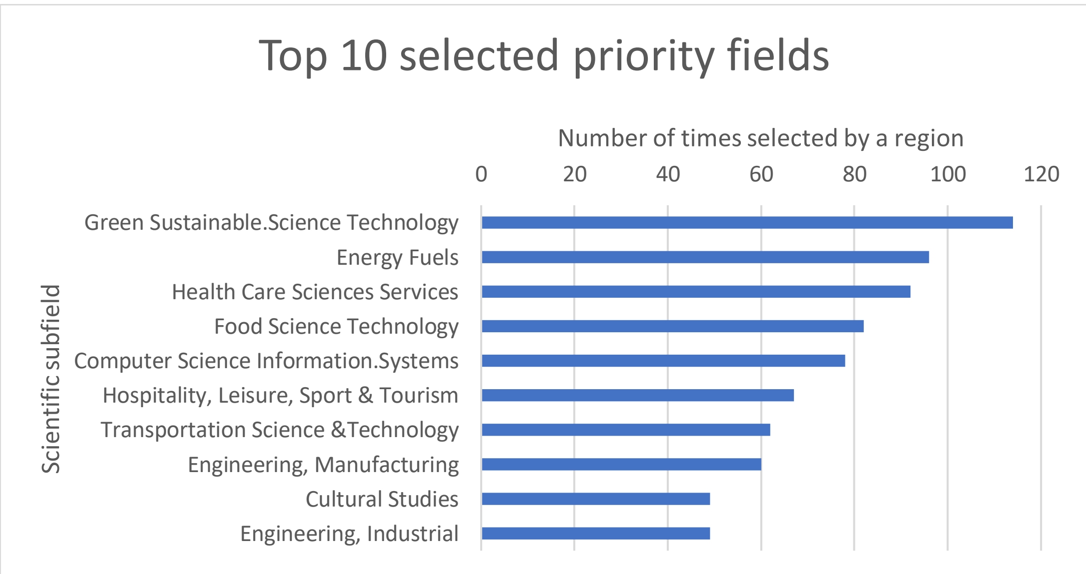


Figure 13: Chart showing the top ten selected scientific subfields in a region's prioritisation portfolio in the RIS3 program.

To give an overview of which priority fields are generally prioritised together, a priority relatedness space is displayed in Figure 14. At first sight it is clear that not much clustering can be found in this space. This indicates that not much similarities can be found in region's prioritisation portfolios, which seems to implicate that regions are indeed able to form unique prioritisation strategies. However, general engineering fields are placed at a distance from social fields, meaning that these fields are typically not found together in the same prioritisation portfolio. Additionally, the popular fields from Figure 13 can be found relatively close together in the centre of the space. This indicates that the popular fields can be found in all sorts of prioritisation strategies and are often prioritised together in the same region. An overview of which regions have similar prioritisation portfolios can be found in Appendix D.

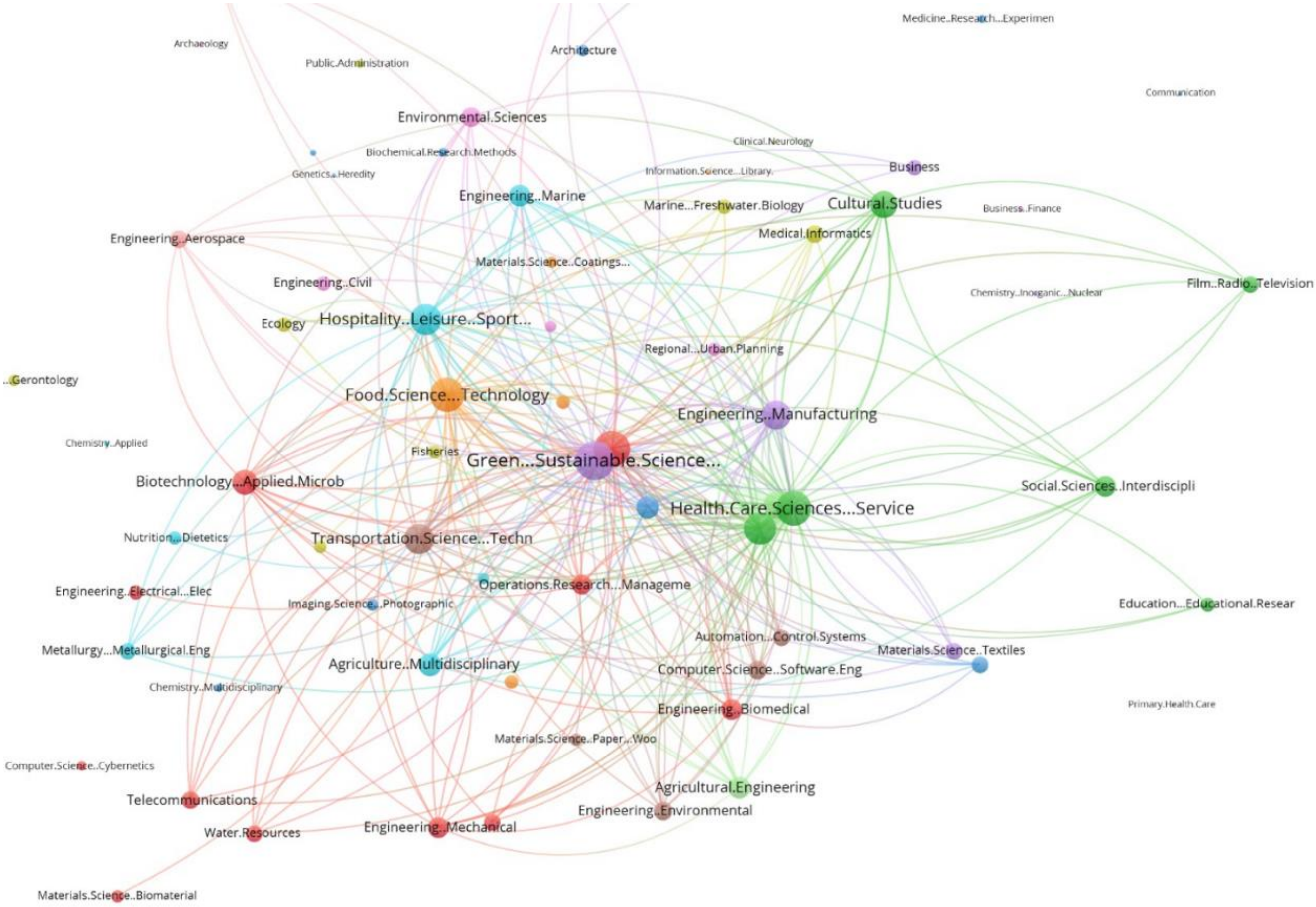


Figure 14: Priority relatedness space showing the relatedness across prioritised scientific subfields based on co-occurrence in European regions (2010-2011).

4.2. Regression analysis

To get a better understanding of the relations that exist between the discussed concepts, multiple linear regression models are constructed, testing the hypotheses. Considering the missing data points on the institutional variables explaining prioritisation strategy, these models are presented in a control model and a full model to provide an additional robustness check. The control model is constructed using original data and the full model is constructed using imputed data from the multiple imputation method. First, the relations between the relatedness density and institutional context with the regional complexity are tested in Model 1, of which the results are listed in Table 4 . Both model types show a high R^2 and a high adjusted R^2 , implicating a high explanatory power.

In the knowledge base model, relatedness density shows a significant positive relation. The control variables population and education level show a significant negative and positive relation respectively. The effect of population is surprising, since metropolitan areas show a higher publication output compared to other regions, but here the opposite is found for knowledge complexity. A potential explanation is that the institutional elements deemed relevant for complex knowledge development, quality of government and institutional thickness, are aided by smaller coherent regions. This is supported by the lack of significance for population in the institutional model, which means that the institutional variables explain the variation inferred by population size, where relatedness density does not. A positive association between the regional education level and knowledge complexity is to be expected, but this is again only found in the knowledge base model. Interestingly, this indicates that education level has a more institutional relation with regional complexity than one based on knowledge capabilities. Education level shows a larger effect on regional complexity than the relatedness density compared to variable size, implicating that education level is a stronger predictor for knowledge complexity than relatedness density. However, relatedness density still shows a noticeable positive relation and therefore hypothesis 1 can be accepted.

In the institutional model, both the institutional thickness and the quality of government show a significant positive relation with knowledge base complexity. Both show higher coefficients compared to variable size and lower significance scores than the control variables, indicating that they explain knowledge base complexity better than the control variables. The quality of government shows a notable larger coefficient than the institutional thickness relative to variable size, which means that quality of government is a stronger predictor for regional complexity than institutional thickness. Nonetheless, both show a notable positive effect and therefore hypotheses 2 and 3 can be accepted.

Table 4: Results of the multiple regression models in Model 1 with regional knowledge base complexity as the dependent variable.

Model 1 (DV regional complexity)		
Model type	Knowledge base	Institutional
Intercept	14,677 (0,000)***	20,032 (0,000)***
Relatedness density	0,002 (0,000)***	
Institutional thickness		0,058 (0,003)**
Quality of Government		0,231 (0,000)***
Population	-1,512E-06 (0,000)***	-1.722E-07 (0,633)
GDP per capita	1,995E-05 (0,604)	-6.833E-05 (0,491)
Education level	0,293 (0,000)***	0,024 (0,792)
R^2	0,482	0,528
Adjusted R^2	0,475	0,505
N	299	109

*Significance codes: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05*

These results confirm that the regional elements, theoretically stated to be important for evolutionary knowledge dynamics, are strongly associated with the level of knowledge development in the region. Furthermore, the similarly high level of explanatory power for the institutional model supports the notion that the institutional context has an equally strong association with the level of knowledge base complexity as the knowledge base characteristics. More specifically, the governmental context and interorganisational density could be just as important in knowledge dynamics and in the smart specialisation program. Ultimately, this raises the question how the knowledge-institutional configuration affects current RIS3 strategies and which knowledge-institutional elements are involved in this aspect. Therefore, the effects of the knowledge-institutional configurations on the prioritisation strategy are tested in Model 2, 3 and 4. Model 2 tests the effects on strategy complexity, Model 3 on strategy relatedness and Model 4 on strategy diversity respectively.

Model 2, displayed in Table 5, shows reasonable R^2 and adjusted R^2 values across the model types. Only the adjusted R^2 of the knowledge base model is a little low with a value under 0,1. In the knowledge base model, the relatedness density shows a significant positive effect on the complexity of the prioritisation strategy, with no significant effects found for the control variables. Therefore, hypothesis 4a can be accepted. This confirms that regions with more low-risk diversification options are better able to integrate complexity in their prioritisation strategy.

In both institutional models HES representation, open exports and GDP per capita show a significant effect on the complexity of the prioritisation strategy. A significant negative effect is found for effective public engagement in the full model, but this is not validated by the control model, which shows a high p-value and is therefore disregarded. Control variable GDP per capita shows negligible coefficients in both institutional models and subsequently the effect for this variable can be considered negligibly small. HES representation shows a large positive effect on strategy complexity in both models and therefore hypothesis 7b can be accepted. Although there is current understanding that the HES has an important role in generating knowledge and regional capabilities, this adds new insights for policy prioritisation. Namely, if regions look to find (related) complexity in their prioritisation strategy, the HES actors are crucial actors to listen to in the prioritisation process.

Open exports show a negative effect on strategy complexity in both models where a positive effect was expected. Hypothesis 8b is therefore rejected. Recent findings agree that this result is indeed odd, as it was found that regions can rely on extra-regional capabilities in order to compensate for missing capabilities (Balland & Boschma, 2020). An obvious explanation could be that regions were not aware of these opportunities or hesitant for the practical implications at the time of prioritisation, as these possibilities are still being researched. A second explanation lies in the chosen indicator of regional open exports. Lagging regions with low economic and knowledge capabilities probably have a relatively small regional market for the basic products that are produced. A large share of production will therefore be exported for extra-regional supply-chains or consumption. For these type of regions, prioritising complexity is not an option, based on basic capability requirements and therefore the regional openness is of little help. Moreover, these inter-regional relations are not as much based on knowledge exchange, but rather on products. This explanation is supported by Figure 16, which shows a high regional openness score for lagging regions.

Effects for public experimentation and effective public engagement could not be found and subsequently hypothesis 5b and 6b are rejected as well.

Table 5: Results of the multiple regression models in Model 2 with the complexity of the prioritisation strategy as the dependent variable.

Model 2 (DV: complexity of the prioritisation strategy)			
Model type	Knowledge base	Institutional (Control model)	Institutional (Full model)
Intercept	26,024 (0,000)***	30,403 (0,005)**	35,646 (0,000)***
Relatedness density	0,419 (0,005)**		
Public experimentation		-684,527 (0,639)	213,274 (0,496)
Effective public engagement		-0,074 (0,639)	-0,223 (0,042)*
HES representation		16,035 (0,013)*	12,100 (0,004)**
Open exports		-16,131 (0,044)*	-10,545 (0,046)*
Population	2,126E-08 (0,956)	1,423E-08 (0,979)	-2,916E-08 (0,943)
GDP per capita	0,000 (0,186)	0,000 (0,011)*	0,000 (0,006)**
Education level	-0,45 (0,679)	0,101 (0,445)	0,176 (0,069)
Institutional decentralization	-0,33 (0,771)	-0,077 (0,591)	-0,001 (0,990)
R^2	0,126	0,224	0,163
***Adjusted* R^2**	0,085	0,127	0,119
N	114	73	163

*Significance codes: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05*

Model 3, displayed in Table 6, shows sufficient R^2 and adjusted R^2-values for all model types, with high values for the institutional control model compared to the full model. For this reason, the institutional control model is taken as leading over the institutional full model in interpreting the effects on strategy relatedness. In the knowledge base model, significant relations are found for the relatedness density and the population. A similar negative effect for population size, as was found for knowledge base complexity, is shown. This hints that, like for the regional complexity, institutional coherency benefits in smaller regions enable these regions to achieve a higher level of strategy relatedness. Relatedness density shows a reasonable positive effect on the relatedness of the prioritisation strategy and therefore hypothesis 4b can be accepted. This indicates that the regions with more diversification opportunities are able to achieve a higher amount of relatedness for their prioritisation strategy to their knowledge base.

The institutional control model shows significant positive effects for effective public engagement and education level. Both show a similar effect size which is reasonable relative to variable proportions. Therefore, hypothesis 6a can be accepted. This implicates that the capability of regional governance to effectively engage and coordinate regional actors enables the region to find higher relatedness in its prioritisation strategy. Kroll (2017) and Benner (2019) suggest that especially for lagging regions this governmental attribute is improved by participation in the RIS3 program. If this is the case, this would imply that these regions improve their ability to reduce risk in prioritisation procedures in future innovation policies. No relations can be inferred for the business representation and subsequently hypothesis 7c is rejected. Education level shows low p-values across the model types, implicating that it forms a predictive indication for the strategy relatedness. This is to be expected, since the more people practice science and knowledge development in the region, the more potential to generate relatedness density and the more awareness about diversification opportunities there is likely to be across actors in the prioritisation process.

Table 6: Results of the multiple regression models in model 3 with the relatedness of the prioritisation strategy as the dependent variable.

Model 3 (DV: relatedness of the prioritisation strategy)			
Model type	Knowledge base	Institutional (Control model)	Institutional (Full model)
Intercept	0,367 (0,000)***	0,439 (0,000)***	0,472 (0,000)***
Relatedness density	0,008 (0,000)***		
Effective public engagement		0,005 (0,002)**	0,002 (0,120)
Private sector representation		0,014 (0,765)	-0,46 (0,230)
Population	-1,390E-08 (0,001)**	-5,429E-09 (0,210)	-1,302E-08 (0,002)**
GDP per capita	9,349E-07 (0,326)	-4,289E-07 (0,683)	2,137E-06 (0,015)*
Education level	0,002 (0,059)	0,004 (0,008)**	0,002 (0,058)
Institutional decentralization	0,001 (0,463)	-0,001 (0,404)	0,001 (0,384)
R^2	0,112	0,434	0,214
***Adjusted* R^2**	0,107	0,384	0,183
N	114	73	163

*Significance codes: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05*

Model 4, displayed in Table 7, shows low statistical power with only the R^2-value of the institutional control model rising above 0,1. Moreover, no statistically significant effects on the diversity of the prioritisation strategy can be found across the model types. Therefore, hypotheses 4c, 5a, 7a and 8a are rejected. It was expected that especially for lagging regions this type of prioritisation strategy would be an opportunity to catch up and improve their regional capabilities (Heimeriks et al., 2019). However, the results do not indicate that these opportunities are currently consciously pursued or unconsciously seized by the involvement of regional actors in the prioritisation process, nor influenced by the subjected regional features. An explanation for this phenomenon could be that it proves difficult for lagging regions to implement diversity in their strategy and achieving an acceptable level of relatedness at the same time, given their relative low relatedness density.

Table 7: Results of the multiple regression models with the diversity of the prioritisation strategy as the dependent variable.

Model 4 (DV: diversity of the prioritisation strategy)			
Model type	Knowledge base	Institutional (Control model)	Institutional (Full model)
Intercept	3,268 (0,000)***	3,558 (0,000)***	3,705 (0,000)***
Relatedness density	0,008 (0,103)		
Public experimentation		10,668 (0,565)	-0,137 (0,991)
HES representation		-0,143 (0,390)	-0,267 (0,104)
Open exports		-0,547 (0,077)	-0,120 (0,552)
Population	2,335E-08 (0,075)	1,130E-08 (0,381)	2,329E-08 (0,133)
Gdp per capita	-2,491E-06 (0,414)	1,109E-06 (0,754)	-2,941E-06 (0,317)
Education level	-0,007 (0,074)	-,002 (0,551)	-0,003 (0,376)
Institutional decentralization	0,003 (0,387)	,004 (0,261)	0,000 (0,968)
R^2	0,088	0,154	0,060
***Adjusted* R^2**	0,045	0,080	0,017
N	112	65	161

*Significance codes: '***' p < 0.001, '**' p < 0.01, '*' p < 0.05*

These results confirm that certain knowledge and institutional features involved in the RIS3 program, guide for some degree the level of complexity and relatedness of a region's prioritisation strategy. Especially the institutional model for strategy relatedness shows high explanatory power, indicating that the institutional features have a high influence on the establishment of the prioritisation strategy. Moreover, it indicates that institutional elements play an important role in the policy domain of regional knowledge development.

4.3. Cluster Analysis

The regression models provided valuable insight in the general knowledge-institutional dynamics across European regions. To get a sense of how the regional features, assessed in the regression models, are distributed in different type of regions, a cluster analysis is conducted.

In the cluster analysis, The regions of which the prioritisation strategies are collected, are first divided in three clusters based on their knowledge-institutional configurations in Figure 15, which is labelled as a region's level of regional development. The expected division between lagging, intermediate and advanced knowledge developing regions is clearly visible in these clusters. Table 8 shows a sufficient amount of regions represented in each cluster. As expected, the advanced cluster includes the least amount of regions of the three, as only a limited amount of regions are expected to possess very high capabilities (McCann & Ortega-Argilés, 2015; Balland & Rigby, 2017). This again shows that the institutional context gives a similar indication of regional development as the state of the knowledge base would.

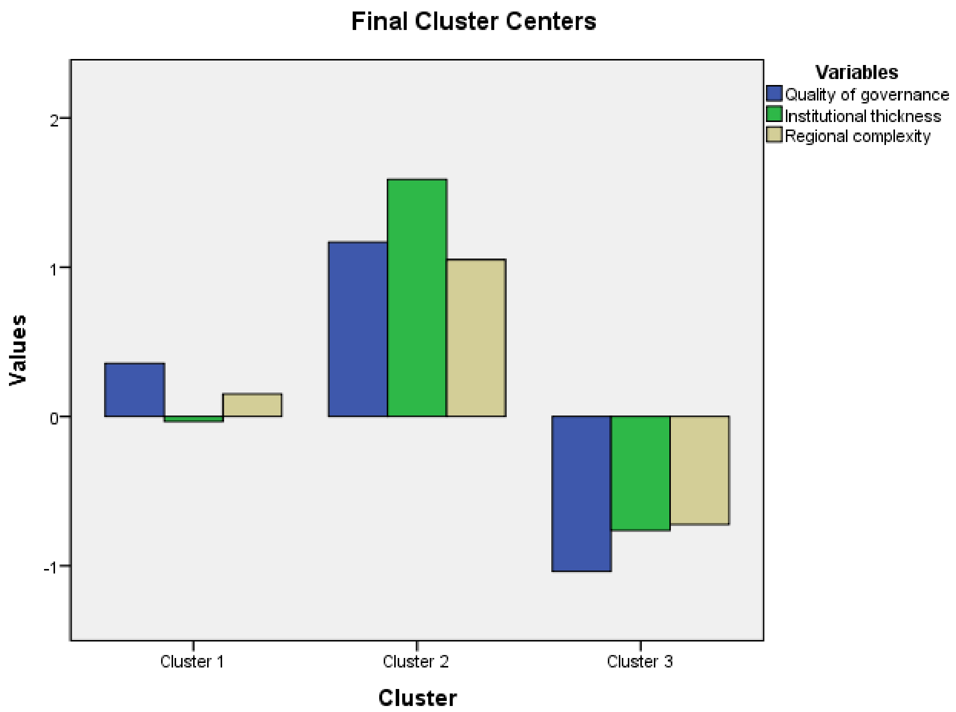


Figure 15: Clustering of European regions based on their knowledge-institutional configurations.

Table 8: Number of regions per cluster in Figure 15

Number of regions per cluster	
Cluster 1 (intermediate)	74
Cluster 2 (advanced)	30
Cluster 3 (lagging)	59
Total	163

An attempt was made to run the regression Models 2, 3 and 4 for each cluster individually to test if the effects of the knowledge-institutional configuration on the prioritisation strategy characteristics differ between types of regions. However, in these models no sufficient statistical power could be reached and therefore no solid effects could be attained.

Nonetheless, to get a sense of how the institutional prioritisation features, discussed in this study, behave in these three different types of regions, a second cluster analysis was performed as shown in Figure 16. Similar scores for the knowledge-institutional attributes from Figure 15 and a similar regional cluster representation, listed in Table 9, in respect to the previous cluster analysis can be found. This confirms that the institutional prioritisation features indeed generally behave according to a region’s knowledge-institutional configuration. This means that these clustering results give a valid representation of the distributions in the advanced, intermediate and lagging regions in terms of regional development. Unsurprisingly, some institutional features show their highest values in the advanced regions, lowest values in the lagging regions and in-between values in the intermediate regions. This is the case for the PP competition, PP control of corruption risks and private sector representation. Therefore, it indicates that elements of effective public engagement and level of private involvement in knowledge development are in line with regional development levels.

Especially interesting are the institutional features that deviate from this pattern. The PP administrative efficiency scores highest in the intermediate regions with quite a large gap to the other two clusters, which seems to influence general effective public engagement as well. Pragmatically this makes sense, since it will probably proof harder for high developed regions, despite superior institutional capabilities, to keep track of all activities going on compared to the more coherent intermediate regions. This is likely to affect the PP administrative efficiency and effective public engagement in general. This gives intermediate regions an institutional edge to prioritise relatedness, which is probably partly compensated by the knowledge capabilities of the advanced regions.

HES representation scores highest in the lagging regions, supporting the notion that the HES has a relatively large responsibility for knowledge development in lagging regions. This result stresses the importance of these actors in the prioritisation process for these regions. It seems that the public experimental context and open exports score in sync, with relative high values in the advanced and lagging regions and lower values for the intermediate regions. The regional openness could be high in advanced regions, because they possess exclusive products that they export to other regions. The regional openness for lagging regions could be high, because these regions rely on economic and knowledge capabilities of other regions to buy or further develop their products. For the public experimental context, one could argue that relative public experimental expenses are high in lagging regions, because of their relatively low regional GDP. However, this should still not explain the observed relative lack of public experimentation in regions that are expected to be more focused on knowledge development. Therefore, these results show that the governance in intermediate regions does not seem confident or dedicated to stimulate and perform experimentation, probably because they rather focus on other aspects of public spending. Advanced regions do seem to rate public experimentation highly, since these regions score highest, despite their high GDP levels and a considerable level of private involvement. This supports the notion raised in Figure 8, that governance experimentation goes hand in hand with private involvement.

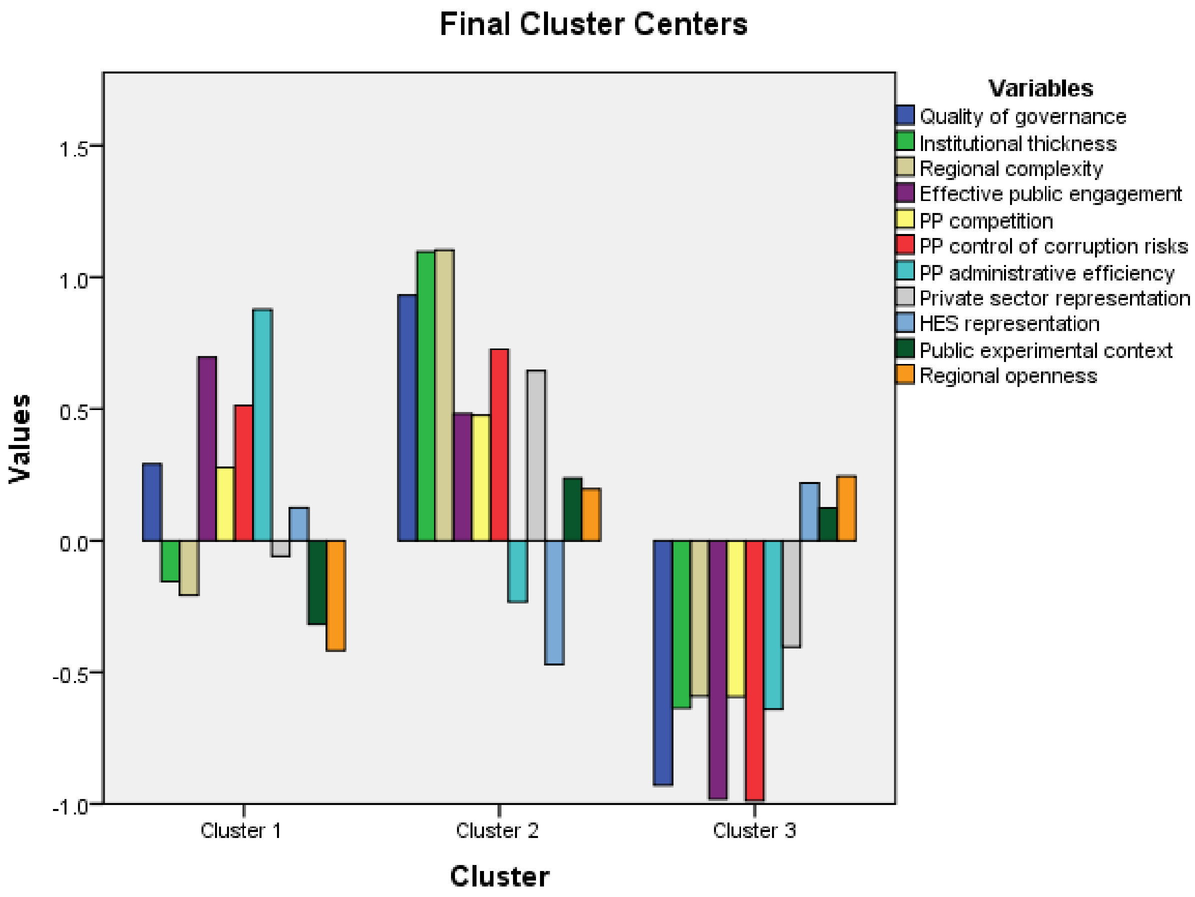


Figure 16: Clustering that shows the distribution of institutional prioritisation features among the knowledge-institutional configurations.

Table 9: Number of regions per cluster in Figure 16

Number of regions per cluster	
Cluster 1 (intermediate)	57
Cluster 2 (advanced)	44
Cluster 3 (lagging)	66
Total	163

The advanced regions show, as expected, the most capabilities to rely on. Following theory and earlier findings, this means that these regions should have an easy time finding appropriate high level priorities. However, some institutional features prove to be high in the lagging and intermediate regions as well. This could mean that these types of regions could use these features to their advantage or at least be aware of their strengths and weaknesses in order to participate in effective knowledge development. Intermediate regions can rely on a good quality of government, a high potential for effective public engagement and a certain level of private involvement. The previous analyses have shown that the overall quality of government can be associated with complex knowledge development and that effective public engagement is beneficial for strategy relatedness. Furthermore, the cluster analysis indicates that steps have to made in regard of the duality of public experimental context and private representation to increase institutional thickness and to get closer to the advanced regions. So, a bigger commitment in public-private experimentation, making use of the governmental capabilities present by taking a leading governance role in pursuing substantiated priorities, should be a step in the right direction for these regions.

The lagging regions lag behind in most features, but do show some assets in the form of HES representation, a decent degree of public experimentation and connectedness with other regions. Given the important role in knowledge development of the HES for these regions, combined with the shown positive effect of HES involvement on strategy complexity, it would seem that basing the priorities on the HES capabilities is a promising strategy. The presence of a certain degree of public experimentation in the lagging regions could be the asset that these regions need to divert to more diverse knowledge development. Public incentives and programs could be designed with this goal in mind to generate more diversification opportunities. Additionally, these regions could try to use their interregional connections to benefit from capabilities of extra-regional partners (Balland & Boschma, 2020). However, such a dynamic has not been substantiated for the prioritisation strategy in this study.

5. Discussion

5.1. Theoretical implications

Based on evolutionary theory and its connection with literature on EU regional innovation policy the dynamics of the regional knowledge-institutional configuration in relation to knowledge development and RIS3 prioritisation were explored. A scientific knowledge perspective was taken in this regard as it provides valuable insight on underlying regional knowledge capabilities, which are stated to form essential building blocks along the learning trajectory (Boschma et al., 2014; Heimeriks et al., 2019).

In assessing the scientific output of regions, patterns were found consistent with previous studies on evolutionary knowledge development (Heimeriks et al., 2019; Balland et al., 2019). As expected, scientific output was found to be spiky and concentrated in metropolitan regions, but relatedness density and complexity showed a different distribution (Florida, 2005). To explain these knowledge development disparities, institutional literature points to quality of governance and institutional thickness as relevant regional elements (Rodriquez Pose & Di Cataldo**,** 2015; Benner, 2019). This study was able to show in one research design that both streams rightfully argue for their involvement in knowledge development dynamics. It was found that the level of relatedness density, quality of government and institutional thickness in a region are associated with the complexity level of its knowledge base. Therefore, upholding the notion that diversification opportunities and a supportive institutional environment aid in traveling along knowledge trajectories. Moreover, indications were found that the knowledge and institutional regional elements account partly for different aspects in explaining variation in knowledge development. This supports the notion that institutional elements account for the found knowledge development disparities. Supplementary to these findings, the control variables indicated that a smaller region size partly covers for the effect of the institutional context. If this indication for institutional coherency is found to be grounded, it could be a counterforce to proximity advantages in the accumulation of knowledge that are typically found in highly populated regions (Heimeriks & Boschma, 2013; Balland & Rigby, 2013).

Additional insight on these matters was provided by constructing relatedness spaces for the knowledge base and institutional context. The knowledge space confirms the theoretical notion that broad knowledge fields are not structured equally and dynamics differ from each other (Asheim & Coenen, 2005). It is therefore important to interpret the results found in this study as general guidelines, which may differ across sectors. A map of the institutional density supported the notion that some regions might compensate or be hampered by their institutional capabilities in achieving knowledge complexity when comparing it to European knowledge distributions (Balland et al., 2019). Additionally, more specific insight on the arrangement of institutional features in European regions was provided, with the institutional relatedness space. It showed that features associated with structure and open creativity are often found in different regions. Interestingly, this finding corresponds with the classification of either having a liberal or coordinated market structure. Therefore, it is consistent with the theoretical notion that the market structure partly shapes the types of institutional features that are developed (Boschma, 2004; Hall & Soskice, 2001).

By consulting literature on smart specialisation dynamics, expectations were formed on how regional knowledge and institutional features might influence the type of prioritisation strategy pursued in the RIS3 program. This was done to connect the acquired understanding on knowledge development dynamics with current policy making and therefore provide useful additions to smart specialisation theory.

Regarding the prioritised topics, interesting insight was found in the high number of priorities in the fields of green technology and health care. Although this seems to be desirable facing current societal challenges, not every region possesses the right capabilities to develop these fields themselves and this could subsequently hamper the smart specialisation concept (Santoalha & Boschma, 2019). Therefore, to align smart and sustainable growth and make effective use of the input from the prioritisation process, it would be more desirable in some cases if sustainable growth would be a reoccurring theme in the priority implementation, instead of an individual priority (Foray, 2016).This could be integrated in the sense that the priorities which fit the regional context and actors best are implemented with an eye for planet and people. Moreover, this would result in more specific priorities than the broad topics of green technology or health care, helpful for achieving the level of granularity that smart specialisation theory is aiming for (Foray et al, 2018). Consistent with theory, however, were the findings that regions form unique prioritisation that seem to be tailored to the specific region (Foray et al., 2011). Regarding the relations with the regional knowledge base, the relatedness density was found to positively influence the relative complexity and relatedness of the prioritisation strategy, as should be expected given the RIS3 guidelines (Boschma & Gianelle, 2014). Therefore, these results suggest that overall actors in the prioritisation process are indeed able to consciously or unconsciously find priorities within appropriate boundaries of risk and gains considering their knowledge base (Balland et al., 2019).

For the institutional context, clear effects of the subjected institutional features were not abandonedly found in this study. It could be that the actors in the prioritisation process are not able to align all expected institutional features in their prioritisation strategy in the current policy design. A second explanation could be that the effect of the institutional context is not reflected clearly in individual institutional features, but rather as a whole, since these features' interplay makes a region's unique institutional context (Rodriquez-Pose, 2013; McCann & Ortega-Argilés, 2014). In that case, the observed effects would point at individual institutional features that are very important in achieving certain strategy attributes. Effects found in agreement with literature indicate a positive influence for HES representation and effective public experimentation for strategy complexity and relatedness respectively. The effect of regional openness showed a negative effect on strategy complexity in contrast to current theoretical understanding, although with the notion that it has proven to be a difficult concept to operationalise (Balland & Boschma, 2020). Additionally, the control variables showed again a role for institutional coherency advantages in the found effect that a smaller population size partly covers for institutional capabilities in relation to the variation of strategy relatedness.

In case of the prioritisation diversity little to none conformity with the knowledge base could be observed. That trend continued for both knowledge and institutional features in the regression model. This seems to be a missed opportunity for smart specialisation policy, since previous studies indicate that prioritising diversity can be a beneficial strategy for regions with lower knowledge-institutional capabilities (Heimeriks et al,, 2019).

In line with the literature, it was found that a limited number of advanced regions have a relative high degree of institutional features to their disposal (McCann & Ortega-Argilés, 2015; Balland & Rigby, 2017). Following theory and previous findings, this indicates that these regions should have an easy time prioritising high quality and low-risk priorities. Consistent with theory, it was found that these capabilities deteriorate going from advanced to intermediate and lagging regions. That being said, intermediate regions still show favourable governance capabilities with which they could try to close the institutional thickness deficit to advanced regions. Despite their institutional shortcomings, lagging regions have no other choice than to make the best use of the few assets they have in the form of HES involvement, public experimentation and interregional market connections. The reasonable amount of public experimentation might provide room for the implementation of more diverse strategies in these regions (Heimeriks et al., 2019). These findings are in line with Kroll (2017), to which he adds that infrastructural improvements still have to be made to translate the improved prioritisation in entrepreneurial practice.

In terms of contributions, this study thus provides evidence for evolutionary knowledge development, where knowledge capabilities form the building blocks and the institutional context the arena in which it takes place. Specific regional attributes are found to be shaping the evolutionary trajectory, being the amount of diversification opportunities, quality of governance and institutional thickness. These showings indicate that the regional context has to be taken into account in a territorial innovation policy design like the RIS3 program. This study shows that some regional features associated with the aforementioned regional attributes already guide for some extent the direction of the prioritisation strategy. These proven features can be focused on in regions that want to focus on a certain strategy characteristic. Moreover, the features which did not show to have an effect now might prove to be useful to integrate in policy design in the future, since quality of government and institutional thickness proves to be relevant for knowledge development. However, fashionable priorities still seem to be overrepresented in the region's prioritisation strategies, which will likely hinder effective learning efforts. Therefore, a thematic approach for socially relevant goals is suggested in which fitting priorities can be developed. Lastly, it is indicated that policy design and its goals should be adjusted based on the level of regional development. Different types of regions are aided by different kind of knowledge development strategies, which scope should be adjusted based on the vastly different capabilities that regions have to their disposal.

5.2. Limitations

This study also has its limitations. Although the choice for a scientific knowledge perspective and the use of publication data was made carefully, it does not encompass the whole story of regional innovation and the role for the prioritisation strategy in it. Many previous studies have looked at patent data for this purpose, which shows practical capabilities and potential for direct societal impact, which this study does not take into account for the prioritisation strategy. Furthermore, there is room for goals aimed at institutional change in the prioritisation strategy, which has not been looked at in this research design. That being said, the RIS3 priorities are currently not formed and documented with this aspect at its core. Therefore, this opportunity is still underused and difficult to research, but might receive more notice in the future as institutional awareness in this respect increases in the academic and policy arena.

Additionally, the construction of the prioritisation strategy with the EYE@RIS3 database comes with its limitations as well. The database itself does not provide a detailed description for every region to the level of detail that would be desired from a smart specialisation perspective. Whether this is due to the priorities set by the regions themselves or to the documentation method upheld by the European Commission could have implications for the interpretation of the results. By linking the priorities from the database to scientific subfields, more specificity and usability was created. However, this comes with the price that this is, although structured, a subjective procedure that could have influenced the results. Furthermore, the type of innovations that regions pursue, radical or incremental, cannot be obtained from this data source, while this can have implications for the type of institutions involved and achievability of the prioritisation strategy.

The operationalisation of an elusive concept like the institutional context gives rise to some difficulties. Although carefully chosen indicators were used, it is unlikely that every indicator is totally comprehensive for the institutional features that they represent. This has its implications on the results, as was for instance likely the case for regional openness. Rodríguez-Pose & Crescenzi (2008) constructed a more sophisticated indicator for regional openness, which indicates the innovation spill-overs between neighbouring regional innovation systems. Although this might have been a methodological improvement, creating such a sophisticated measure for regional openness was not feasible within the boundaries of this research.

5.3. Practical implications

From a societal perspective, the RIS3 program set the goals of smart, sustainable and inclusive growth (European Commission, 2012). To more effectively approach these goals this study provides practical insight on which contextual features guide the prioritisation process. Moreover, it provides suggestions on which contextual features different types of regions can rely upon and which types of strategies can be pursued under which circumstances.

In terms of smart growth, policy makers generally show the ability to set priorities on a fitting level of complexity and relatedness according to regional capabilities. To further improve RIS3 smartness, policy makers are advised to focus on effective public engagement in order to reduce their knowledge deficit and improve the policy connectedness to the regional context (Foray et al., 2018). Furthermore, when looking at which fields could be selected to bring regional knowledge capabilities to a higher level, policy makers are advised to consult the HES as these actors seem to have the best overview of possibilities and capabilities to act on them. The results provide an interesting perspective on the sustainability of the prioritisation strategy. If the priorities will provide long term growth remains to be seen, however, this societal goal does seem to translate in a high amount of priorities in green tech and health. Policy makers should be careful not to interfere with the smart growth when prioritising such fields and could in such cases be better off by opting for a thematic approach.

From the study's, perspective the interregional inclusiveness can be assessed. Apart from probable improvements in the effective public engagement, lagging regions do not show to have notable gains from the RIS3 program compared to other regions. More capable regions show a larger ability to prioritise related and complex fields, which leaves little chance for the lagging regions to catch up. However, lagging regions do have the opportunity to diversify their knowledge base by prioritising diversely and therefore improving their regional capabilities. Since it is not realistic for these regions to prioritise complexity as more advanced regions do, policy makers should focus public funding and R&D efforts on stimulating diversity and improving institutional features They could do so by paying attention to where higher education sees opportunities and trying to make use of relevant capabilities in regions with which they have close market relations. To improve on the institutional features mentioned in this study, new practices could be integrated in the RIS3 program, as has already been shown to be beneficial for the effective public engagement (Benner, 2019). In intermediate regions, policy makers should take a proactive attitude in knowledge development to enable their decent regional knowledge and institutional features to travel further along the desired trajectories, which in turn supports the achievement of societal goals. Taking a proactive and leading role seems particularly achievable given the relative high amount of governance capabilities in these regions. This way, all types of regions are given an appropriate approach to improve regional capabilities, which would make the RIS3 program more inclusive for all regions.

5.4. Future research

The promising results for assessing knowledge development in relation to knowledge-institutional configurations begs for more research in this area. This might encourage future studies to incorporate both elements when looking at regional development to obtain a more holistic understanding. It would be interesting to delve in more detail and see where knowledge and institutional support overlap and in what aspect they complement each other over time. At least, it should be a call for more attention for the institutional context in regional development and smart specialisation studies, as these regional elements are currently still underrepresented in literature (Rodriquez-Pose, 2020). Future studies could specifically investigate the role of hard versus soft institutions in knowledge development and investigate how regions can obtain these capabilities, as this is still very unclear (Rodriquez-Pose, 2020). Furthermore, the new concept of institutional density should be further developed to get a similar indication for institutional relatedness as for knowledge relatedness. These measures can then be used by policy makers to select promising prioritisation strategies, complementary to their regional knowledge-institutional configuration.

On the topic of smart specialisation, future research should inquire into the dynamics inside the prioritisation process. This includes which type of actors are present, who is leading the conversation and on which regional aspects is focussed as this might very well influence the outcome of the prioritisation strategy. Additionally, where this research focusses on the prioritisation phase it would be a relevant follow-up research to look into how knowledge and institutional elements influence the implementation stage of smart specialisation. Moreover, when sufficient implementation time has passed, future research should assess the prioritisation strategy success rate based on strategy type and type of regional-institutional configuration and compare it to the results found in this research to make an assessment about the RIS3 program.

Future studies analysing similar dynamics but with different scopes, for instance the individual region or sectoral level, can also draw some direction from this research. Although lagging regions show some potential to improve in the RIS3 program, it should be investigated if these regions are practically able to utilise these opportunities. To a lesser extent this is also the case for intermediate regions and therefore both are interesting region types to focus on. Lastly, it is an established notion that knowledge trajectories differ across different sectors (Asheim & Coenen, 2005). One would therefore expect that a similar knowledge-institutional configuration analysis could be performed focusing on the sectoral level, which could be investigated in the future.

6. Conclusion

The study's goal was to understand how the knowledge-institutional configurations of regions relate to knowledge development and how this influences a region's prioritisation strategy. It attempted to do so by first understanding how knowledge and institutional configurations behave in terms of general knowledge development. Secondly, it attempted to find features of these configurations that could affect the prioritisation strategy given the structure of the prioritisation process and investigated how these features are distributed across different levels of regional development.

Hypotheses 1, 2 and 3 could be accepted, which meant that the expected regional knowledge-institutional configurations in relation to knowledge development were found. They confirmed that the relatedness density, quality of government and institutional thickness are associated with the level of a regional complexity. Four effects of the knowledge-institutional context on the prioritisation strategy could be found in the form of hypotheses 4a, 4b, 6a and 7b. The first two confirmed a positive effect of the relatedness density on the complexity and relatedness of the prioritisation strategy. The last two confirmed a positive effect of PP governance on strategy relatedness and a positive effect of HES representation on strategy complexity. Additionally, a negative effect for regional openness was found for strategy complexity where a positive effect was expected.

These findings implicate that the institutional elements of quality of government and institutional thickness relate just as well to the complexity of regional knowledge development as the regional diversification opportunities do. The conclusive findings on the RIS3 prioritisation strategies indicate that regions are generally capable of prioritising according to the diversification opportunities, based on their regional knowledge base. Furthermore, the representation of HES regional actors proves to result in more ambitious prioritisation and the capability of the regional governance to connect with regional actors aids in finding related strategies. In general the explanatory power of the institutional features were higher than that of the knowledge capabilities, indicating that RIS3 prioritisation is more guided by the institutional context than the knowledge base.

Further findings suggest that regions tend to prioritise popular societal topics such as climate change and health care, despite being unrelated to the knowledge base, which may compromise the effectiveness of the smart specialisation rationale. Additionally, the strategy option to further diversify a region's knowledge development is currently not pursued, while this may be a promising strategy for lagging regions. Considering the different levels of regional development, it was unsurprisingly found that leading regions in regional development score highest in most institutional features deemed relevant for a prosperous strategy design. However, intermediate regions show favourable governmental capabilities that could be mobilised to close the institutional thickness deficit to the advanced regions and presumably improve their knowledge development. Lagging regions display some public experimental activity that could be utilised to improve the diversity of knowledge development. Additionally, they show a high dependence on the HES and these actors should therefore be fundamentally engaged in their prioritisation process .

In general it can be concluded that knowledge and institutional elements play an important role in knowledge development that partly resonates through into policy design. In recognizing how these elements differ between regions, innovation strategies can be more carefully aligned with the knowledge-institutional configuration in its design and in the goals that it pursues.

7. References

Adams, J. D. (1990). Fundamental stocks of knowledge and productivity growth. Journal of political economy, 98(4), 673-702.

Arthur, W. B. (2009). The nature of technology: What it is and how it evolves. Simon and Schuster.

Asheim, B. T., & Coenen, L. (2005). Knowledge bases and regional innovation systems: Comparing Nordic clusters. Research policy, 34(8), 1173-1190.

Balland, P. A. (2017). Economic Geography in R: Introduction to the EconGeo package. Available at SSRN 2962146.

Balland, P. A., & Rigby, D. (2017). The geography of complex knowledge. Economic Geography, 93(1), 1-23.

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

Balland, P. A., & Boschma, R. (2020). Complementary Inter-Regional Linkages and Smart Specialization: an Empirical Study on European Regions (No. 2023). Utrecht University, Department of Human Geography and Spatial Planning, Group Economic Geography.

Benner, M. (2014). From smart specialisation to smart experimentation. Zeitschrift für Wirtschaftsgeographie, 58(1), 33-49.

Benner, M. (2019). Smart specialization and institutional context: the role of institutional discovery, change and leapfrogging. European Planning Studies, 27(9), 1791-1810.

Boschma, R. (2004). Competitiveness of regions from an evolutionary perspective. Regional studies, 38(9), 1001-1014.

Boschma, R. (2017). Relatedness as driver of regional diversification: A research agenda. Regional Studies, 51(3), 351-364.

Boschma, R., & Gianelle, C. (2014). Regional branching and smart specialization policy. JRC technical reports, (06/2104).

Boschma, R., Heimeriks, G., & Balland, P. A. (2014). Scientific knowledge dynamics and relatedness in biotech cities. Research Policy, 43(1), 107-114.

Boschma, R., Coenen, L., Frenken, K., & Truffer, B. (2017). Towards a theory of regional diversification: combining insights from Evolutionary Economic Geography and Transition Studies. Regional studies, 51(1), 31-45.

Bryman, A. (2012). Social Research Methods (4th ed.). Retrieved from https://www.academia.edu/30520568/Social_Research_Methods_4th_Edition_by_Alan_Bryman.pdf

Cantner, U., Meder, A., & Ter Wal, A. L. (2010). Innovator networks and regional knowledge base. Technovation, 30(9-10), 496-507.

Carayannis, E., & Grigoroudis, E. (2016). Quadruple innovation helix and smart specialization: Knowledge production and national competitiveness. Форсайт, 10(1 (eng)).

Cardinal, L. B., Alessandri, T. M., & Turner, S. F. (2001). Knowledge codifiability, resources, and science-based innovation. Journal of knowledge management, 5(2), 195-204.

Cebolla, R. O., & Navas, C. (2019). Supporting hydrogen technologies deployment in EU regions and member states: the smart specialisation platform on energy (S3PEnergy). International Journal of Hydrogen Energy, 44(35), 19067-19079.

Charron, N., Dijkstra, L., & Lapuente, V. (2014). Regional governance matters: Quality of government within European Union member states. Regional Studies, 48(1), 68-90.

Cooke, Phil, Björn Asheim, Jan Annerstedt, Jiri Blazek, Ron Boschma, Danes Brzica, Åsa Lindholm Dahlstrand et al. "Constructing regional advantage: Principles, perspectives, policies." (2006).

Cooke, P., & Leydesdorff, L. (2006). Regional development in the knowledge-based economy: The construction of advantage. The journal of technology Transfer, 31(1), 5-15.

De Noni, I., Ganzaroli, A., & Orsi, L. (2017). The impact of intra-and inter-regional knowledge collaboration and technological variety on the knowledge productivity of European regions. Technological Forecasting and Social Change, 117, 108-118.

Dosi, G. (1982). Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. Research policy, 11(3), 147-162.

Dosi, G., & Nelson, R. R. (2013). The Evolution of Technologies: An Assesment of the State-of-the-Art. Eurasian Business Review, 3(1), 3–46. https://doi.org/10.14208/BF03353816

van Eck, N. J., & Waltman, L. (2013). VOSviewer manual. Leiden: Univeristeit Leiden, 1(1), 1-53.

Etzkowitz, H., & Leydesdorff, L. (1995). The Triple Helix--University-industry-government relations: A laboratory for knowledge based economic development. EASST review, 14(1), 14-19.

European Commission. (2012). Guide to Research and Innovation Strategies for Smart Specialisations (RIS 3). Retrieved on 3-12-2019 from https://ec.europa.eu/jrc/sites/jrcsh/files/RIS3_GUIDE_FINAL.pdf

European Commission. (2018a). Eye@RIS3: Innovation Priorities in Europe. Retrieved on 16-12-2019, from https://s3platform.jrc.ec.europa.eu/map?p_p_id=captargmap_WAR_CapTargMapportlet&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-1&p_p_col_count=1

European Commission. (2018b). Higher Education for Smart Specialisation. Retrieved from https://s3platform.jrc.ec.europa.eu/documents/20182/222215/HESS+Handbook+_Online.pdf/775c8e1f-dd57-4ab0-83a1-0547d8a3e908

Eurostat. (2007). Europa - RAMON - Classification Detail List. Retrieved 15-01-2020, from https://ec.europa.eu/eurostat/ramon/nomenclatures/index.cfm?TargetUrl=LST_NOM_DTL&StrNom=CL_NABS07&StrLanguageCode=EN&IntPcKey=40311324&StrLayoutCode=HIERARCHIC

Fazekas, M. I. H. Á. L. Y. (2017). Assessing the quality of government at the regional level using public procurement data. Digiwhist. http://digiwhist. eu/publications/quality-ofgovernment.

Florida, R. (2005). THE WORLD IS SPIKY Globalization has changed the economic playing field, but hasn't leveled it. Atlantic monthly, 296(3), 48.

Fraley, C., & Raftery, A. E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. The computer journal, 41(8), 578-588.

Foray, D. (2016). On the policy space of smart specialization strategies. European Planning Studies, 24(8), 1428-1437.

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.

Foray, D., David, P. A., & Hall, B. H. (2011). Smart specialisation from academic idea to political instrument, the surprising career of a concept and the difficulties involved in its implementation (No. REP_WORK). EPFL.

Foray, D., Morgan, K., & Radosevic, S. (2018). The role of smart specialization in the EU research and innovation policy landscape. European Commission.[online] Available at: http://ec. europa. eu/regional_policy/sources/docgener/brochure/smart/role_smartspecialisation_ri. pdf [Accessed 28 Nov. 2018].

Davidson Frame, J., Narin, F., & Carpenter, M. P. (1977). The distribution of world science. Social studies of science, 7(4), 501-516.

Gianelle, C., Guzzo, F., & Mieszkowski, K. (2018). Smart Specialisation at work: Assessing investment priorities. Retrieved on 22-11-2019 from https://s3platform.jrc.ec.europa.eu/documents/20182/201464/Smart+Specialisation+at+work_Assessing+investment+priorities/a48dd31a-172d-464f-817b-fd811745c9c3

Gianelle, C., Guzzo, F., & Marinelli, E. (2019). SMART SPECIALISATION EVALUATION: SETTING THE SCENE. Retrieved from https://ec.europa.eu/jrc/sites/jrcsh/files/jrc116110.pdf

Glückler, J., & Bathelt, H. (2017). Institutional context and innovation. In H. Bathelt, P. Cohendet, S. Henn, & L. Simon (Eds.), The Elgar companion to innovation and knowledge creation (pp. 121–137). Cheltenham, Northampton: Elgar.

Guevara, M. R., Hartmann, D., & Mendoza, M. (2016). diverse: an R Package to Analyze Diversity in Complex Systems. R J., 8(2), 60.

Hall, P. A., & Soskice, D. (2001). An introduction to varieties of capitalism. op. cit, 21-27.

Heimeriks, G., & Balland, P. A. (2015). How smart is specialisation? An analysis of specialisation patterns in knowledge production. Science and Public Policy, 43(4), 562-574.

Heimeriks, G., & Boschma, R. (2013). The path-and place-dependent nature of scientific knowledge production in biotech 1986–2008. Journal of Economic Geography, 14(2), 339-364.

Heimeriks, G., Li, D., Lamers, W., Meijer, I., & Yegros, A. (2019). Scientific knowledge production in European regions: patterns of growth, diversity and complexity. European Planning Studies, 27(11), 2123-2143.

van der Hel, S. (2016). New science for global sustainability? The institutionalisation of knowledge co-production in Future Earth. Environmental science & policy, 61, 165-175.

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for research metrics. Nature News, 520(7548), 429.

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

Hidalgo, C. A., & Hausmann, R. (2009). The building blocks of economic complexity. Proceedings of the national academy of sciences, 106(26), 10570-10575.

IBM. (2020). Method (Multiple Imputation). Retrieved on 21-04-2020 from https://www.ibm.com/support/knowledgecenter/SSLVMB_24.0.0/spss/mva/idh_idd_mi_method.html

Kates, R. W., Clark, W. C., Corell, R., Hall, J. M., Jaeger, C. C., Lowe, I., ... & Faucheux, S. (2001). Sustainability science. Science, 292(5517), 641-642.

Kleinknecht, A., Van Montfort, K., & Brouwer, E. (2002). The non-trivial choice between innovation indicators. Economics of Innovation and new technology, 11(2), 109-121.

Kroll, H. (2015). Efforts to implement smart specialization in practice—leading unlike horses to the water. European Planning Studies, 23(10), 2079-2098.

Kroll, H. (2017). The challenge of smart specialisation in less favoured regions (No. R1/2017). Arbeitspapiere Unternehmen und Region.

Lagendijk, A., & Cornford, J. (2000). Regional institutions and knowledge–tracking new forms of regional development policy. Geoforum, 31(2), 209-218.

Lambooy, J. G., & Boschma, R. A. (2001). Evolutionary economics and regional policy. The Annals of Regional Science, 35(1), 113-131.

Landabaso, M. (2014). Guest editorial on research and innovation strategies for smart specialisation in Europe. European Journal of Innovation Management.

Landabaso, M., & Foray, D. (2014). From smart specialisation to smart specialisation policy. European Journal of Innovation Management.

Landabaso, M., McCann, P., & Ortega-Argilés, R. (2014). Smart specialisation in European regions: Issues of strategy, institutions and implementation. European Journal of Innovation Management.

Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. Journal of Informetrics, 5(1), 87-100.

Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. Pattern recognition, 36(2), 451-461.

Maskell, P., & Malmberg, A. (1999). Localised learning and industrial competitiveness. Cambridge journal of economics, 23(2), 167-185.

Mazzucato, M. (2011). The entrepreneurial state. Soundings, 49(49), 131-142.

McCann, P., & Ortega-Argilés, R. (2013). Modern regional innovation policy. Cambridge Journal of Regions, Economy and Society, 6(2), 187-216.

McCann, P., & Ortega-Argilés, R. (2014). Smart specialisation in European regions: Issues of strategy, institutions and implementation. European Journal of Innovation Management, 17(4), 409-427.

McCann, P., & Ortega-Argilés, R. (2015). Smart specialization, regional growth and applications to European Union cohesion policy. Regional Studies, 49(8), 1291-1302.

Morgan, K. (2017). Nurturing novelty: Regional innovation policy in the age of smart specialisation. Environment and Planning C: Politics and Space, 35(4), 569-583.

Navarro, M., Gibaja, J. J., Franco, S., Murciego, A., Gianelle, C., Hegyi, F. B., & Kleibrink, A. (2014). Regional benchmarking in the smart specialisation process: Identification of reference regions based on structural similarity. Institute for Prospective and Technological Studies, Joint Research Centre.

Nelson, R. R. (1959). The simple economics of basic scientific research. Journal of political economy, 67(3), 297-306.

Nelson, R. R., & Romer, P. M. (1996). Science, economic growth, and public policy. Challenge, 39(1), 9-21.

OECD. (2013). Innovation-driven Growth in Regions: The Role of Smart Specialisation. Retrieved on 2-12-2019, from https://www.oecd.org/sti/inno/smart-specialisation.pdf

Powell, W. W., & Snellman, K. (2004). The knowledge economy. Annu. Rev. Sociol., 30, 199-220.

Rao, C. R. (1982). Diversity: its measurement, decomposition apportionment and analysis.

Rodríguez-Pose, A. (2013). Do institutions matter for regional development?. Regional studies, 47(7), 1034-1047.

Rodríguez-Pose, A. (2020). Institutions and the fortunes of territories. Regional Science Policy & Practice.

Rodríguez-Pose, A., & Crescenzi, R. (2008). Research and development, spillovers, innovation systems, and the genesis of regional growth in Europe. Regional studies, 42(1), 51-67.

Rodríguez-Pose, A., Di Cataldo, M., & Rainoldi, A. (2014). The role of government institutions for smart specialisation and regional development. S3 Policy Brief Series, 4.

Rodríguez-Pose, A., & Di Cataldo, M. (2015). Quality of government and innovative performance in the regions of Europe. Journal of Economic Geography, 15(4), 673-706.

Rodríguez-Pose, A., & Wilkie, C. (2017). Institutions and the entrepreneurial discovery process for smart specialization. Governing smart specialisation, 34-48.

Rubin, D. B. (1996). Multiple imputation after 18+ years. Journal of the American statistical Association, 91(434), 473-489.

Rubinfeld, D. L. (2000). Reference guide on multiple regression. Reference manual on scientific evidence, 179, 425-469.

Sahal, D. (1981). Alternative conceptions of technology. Research policy, 10(1), 2-24.

Santoalha, A., & Boschma, R. (2019). Papers in Evolutionary Economic Geography# 19.22.

Schumpeter J. (1943) Socialism, Capitalism and Democracy . London : Allen and Unwin .

Shiroyama, H., Yarime, M., Matsuo, M., Schroeder, H., Scholz, R., & Ulrich, A. E. (2012). Governance for sustainability: knowledge integration and multi-actor dimensions in risk management. Sustainability Science, 7(1), 45-55.

Sotarauta, M. (2018). Smart specialization and place leadership: dreaming about shared visions, falling into policy traps?. Regional Studies, Regional Science, 5(1), 190-203.

Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. Journal of the Royal Society Interface, 4(15), 707-719.

Tavassoli, S., & Carbonara, N. (2014). The role of knowledge variety and intensity for regional innovation. Small Business Economics, 43(2), 493-509.

Trippl, M., Zukauskaite, E., & Healy, A. (2019). Shaping smart specialization: the role of place-specific factors in advanced, intermediate and less-developed European regions. Regional Studies, 1-13.

Vedula, S., York, J. G., & Corbett, A. C. (2019). Through the looking-glass: The impact of regional institutional logics and knowledge pool characteristics on opportunity recognition and market entry. Journal of Management Studies, 56(7), 1414-1451.

Waltman, L., Van Eck, N. J., & Noyons, E. C. (2010). A unified approach to mapping and clustering of bibliometric networks. Journal of Informetrics, 4(4), 629-635.

Web of Science. (2018). Research Areas (Categories / Classification). Retrieved 15-01-2020, from http://images.webofknowledge.com/WOKRS534DR1/help/WOS/hp_research_areas_easca.html

8. Appendices

Appendix A: Overview scientific fields

In Table 10 the scientific subfields are listed with their corresponding complexity score in alphabetical order.

Table 10: List of the scientific subfields included in the study with their respective complexity score

Scientific subfield	Complexity	Scientific subfield	Complexity
Acoustics	32,84	Literature	72,13
Agricultural Economics & Policy	29,79	Literature, African, Australian, Canadian	76,93
Agricultural Engineering	21,63	Literature, American	62,62
Agriculture, Dairy & Animal Science	17,09	Literature, British Isles	68,83
Agriculture, Multidisciplinary	16,16	Literature, German, Dutch, Scandinavian	55,13
Agronomy	18,49	Literature, Romance	49,72
Allergy	17,76	Literature, Slavic	2,94
Anatomy & Morphology	18,84	Logic	20,45
Andrology	21,69	Management	57,81
Anesthesiology	54,95	Marine & Freshwater Biology	38,17
Anthropology	49,77	Materials Science, Biomaterials	31,10
Archaeology	51,70	Materials Science, Ceramics	6,98
Architecture	57,51	Materials Science, Characterization & Testing	11,82
Area Studies	57,46	Materials Science, Coatings & Films	8,23
Art	64,73	Materials Science, Composites	14,44
Asian Studies	61,53	Materials Science, Multidisciplinary	0,00
Astronomy & Astrophysics	27,78	Materials Science, Paper & Wood	13,95
Audiology & Speech-Language Pathology	56,57	Materials Science, Textiles	4,52
Automation & Control Systems	20,16	Mathematical & Computational Biology	52,16
Behavioral Sciences	62,61	Mathematics	1,80
Biochemical Research Methods	32,76	Mathematics, Applied	2,84
Biochemistry & Molecular Biology	22,31	Mathematics, Interdisciplinary Applications	14,17
Biodiversity Conservation	50,84	Mechanics	6,12
Biology	32,08	Medical Ethics	59,62
Biophysics	20,12	Medical Informatics	52,66
Biotechnology & Applied Microbiology	16,00	Medical Laboratory Technology	25,65
Business	50,80	Medicine, General & Internal	51,14
Business, Finance	65,38	Medicine, Legal	38,63
Cardiac & Cardiovascular Systems	32,47	Medicine, Research & Experimental	20,19
Cell & Tissue Engineering	38,14	Medieval & Renaissance Studies	70,30

Cell Biology	32,62	Metallurgy & Metallurgical Engineering	6,02
Chemistry, Analytical	14,83	Meteorology & Atmospheric Sciences	38,31
Chemistry, Applied	7,78	Microbiology	37,63
Chemistry, Inorganic & Nuclear	4,72	Microscopy	20,22
Chemistry, Medicinal	10,13	Mineralogy	6,15
Chemistry, Multidisciplinary	7,70	Mining & Mineral Processing	0,20
Chemistry, Organic	13,12	Multidisciplinary Sciences	49,20
Chemistry, Physical	1,37	Music	60,19
Classics	53,18	Mycology	21,08
Clinical Neurology	42,32	Nanoscience & Nanotechnology	10,02
Communication	76,28	Neuroimaging	52,42
Computer Science, Artificial Intelligence	20,58	Neurosciences	46,39
Computer Science, Cybernetics	28,87	Nuclear Science & Technology	12,92
Computer Science, Hardware & Architecture	42,94	Nursing	73,65
Computer Science, Information Systems	24,00	Nutrition & Dietetics	46,58
Computer Science, Interdisciplinary Applications	19,84	Obstetrics & Gynecology	45,76
Computer Science, Software Engineering	29,88	Oceanography	48,69
Computer Science, Theory & Methods	19,80	Oncology	38,91
Construction & Building Technology	11,88	Operations Research & Management Science	31,44
Criminology & Penology	92,24	Ophthalmology	49,71
Critical Care Medicine	52,76	Optics	22,81
Crystallography	4,54	Ornithology	36,69
Cultural Studies	77,19	Orthopedics	46,69
Dance	73,63	Otorhinolaryngology	43,19
Demography	77,52	Paleontology	37,16
Dentistry, Oral Surgery & Medicine	47,34	Parasitology	36,52
Dermatology	28,66	Pathology	29,91
Development Studies	78,16	Pediatrics	46,33
Developmental Biology	40,06	Peripheral Vascular Disease	31,86
Ecology	46,41	Pharmacology & Pharmacy	25,73
Economics	57,63	Philosophy	56,11
Education & Educational Research	45,92	Physics, Applied	7,58
Education, Scientific Disciplines	45,69	Physics, Atomic, Molecular & Chemical	9,63
Education, Special	73,95	Physics, Condensed Matter	1,96
Electrochemistry	8,55	Physics, Fluids & Plasmas	19,20
Emergency Medicine	45,63	Physics, Mathematical	6,23
Endocrinology & Metabolism	32,20	Physics, Multidisciplinary	2,47
Energy & Fuels	22,50	Physics, Nuclear	5,67

Engineering, Aerospace	46,15	Physics, Particles & Fields	13,06
Engineering, Biomedical	30,03	Physiology	39,76
Engineering, Chemical	5,02	Planning & Development	48,26
Engineering, Civil	17,49	Plant Sciences	13,80
Engineering, Electrical & Electronic	20,27	Poetry	100,00
Engineering, Environmental	23,14	Political Science	81,16
Engineering, Geological	21,25	Polymer Science	2,82
Engineering, Industrial	35,05	Primary Health Care	78,66
Engineering, Manufacturing	22,17	Psychiatry	54,66
Engineering, Marine	44,33	Psychology	67,26
Engineering, Mechanical	11,30	Psychology, Applied	64,52
Engineering, Multidisciplinary	12,60	Psychology, Biological	49,49
Engineering, Ocean	43,93	Psychology, Clinical	75,47
Engineering, Petroleum	22,69	Psychology, Developmental	73,57
Entomology	16,27	Psychology, Educational	64,66
Environmental Sciences	24,01	Psychology, Experimental	69,98
Environmental Studies	69,44	Psychology, Mathematical	61,01
Ergonomics	54,28	Psychology, Multidisciplinary	64,12
Ethics	70,68	Psychology, Psychoanalysis	36,69
Ethnic Studies	91,61	Psychology, Social	65,98
Evolutionary Biology	56,28	Public Administration	77,86
Family Studies	81,66	Public, Environmental & Occupational Health	73,22
Film, Radio, Television	81,50	Quantum Science & Technology	18,92
Fisheries	35,37	Radiology, Nuclear Medicine & Medical Imaging	33,51
Folklore	37,43	Regional & Urban Planning	66,75
Food Science & Technology	10,49	Rehabilitation	71,09
Forestry	22,10	Religion	85,60
Gastroenterology & Hepatology	34,23	Remote Sensing	32,65
Genetics & Heredity	49,35	Reproductive Biology	32,93
Geochemistry & Geophysics	30,62	Respiratory System	57,82
Geography	72,14	Rheumatology	47,81
Geography, Physical	48,81	Robotics	23,69
Geology	24,72	Social Issues	88,53
Geosciences, Multidisciplinary	34,74	Social Sciences, Biomedical	93,17
Geriatrics & Gerontology	58,50	Social Sciences, Interdisciplinary	73,50
Gerontology	65,79	Social Sciences, Mathematical Methods	62,08
Green & Sustainable Science & Technology	29,52	Social Work	81,74
Health Care Sciences & Services	83,74	Sociology	70,91
Health Policy & Services	73,13	Soil Science	26,78
Hematology	29,31	Spectroscopy	11,59
History	78,53	Sport Sciences	56,31
History & Philosophy Of Science	79,84	Statistics & Probability	37,39
History Of Social Sciences	85,57	Substance Abuse	67,91

Horticulture	11,78	Surgery	42,60
Hospitality, Leisure, Sport & Tourism	62,86	Telecommunications	41,42
Humanities, Multidisciplinary	74,02	Theater	84,57
Imaging Science & Photographic Technology	36,53	Thermodynamics	0,34
Immunology	37,71	Toxicology	26,38
Industrial Relations & Labor	81,83	Transplantation	20,19
Infectious Diseases	56,89	Transportation	61,09
Information Science & Library Science	65,59	Transportation Science & Technology	31,40
Instruments & Instrumentation	12,23	Tropical Medicine	52,17
Integrative & Complementary Medicine	21,92	Urban Studies	62,69
International Relations	77,91	Urology & Nephrology	30,74
Language & Linguistics	60,71	Veterinary Sciences	31,17
Law	89,07	Virology	40,82
Limnology	37,51	Water Resources	25,63
Linguistics	61,61	Women's Studies	76,73
Literary Reviews	51,36	Zoology	30,14
Literary Theory & Criticism	46,11		

Appendix B: Priorities coding

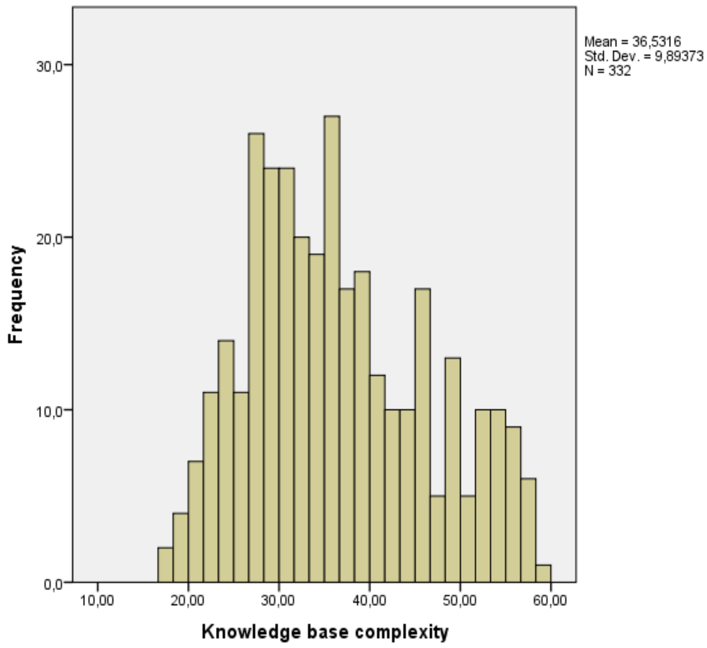
In Table 11 the often appearing terms in the priority description are documented for which the coding proceedings are not completely straightforward. Next to each term the corresponding coded fields are listed in the scientific subfield coding procedure. A "+" means that a fixed combination of two scientific subfields was coded when this priority term was used in the priority description. A "/" means that either one or more of these fields were coded dependent on the further context in the priority description. In a limited number of specific cases, deviations were made from these guidelines to optimise the quality of the coding process.

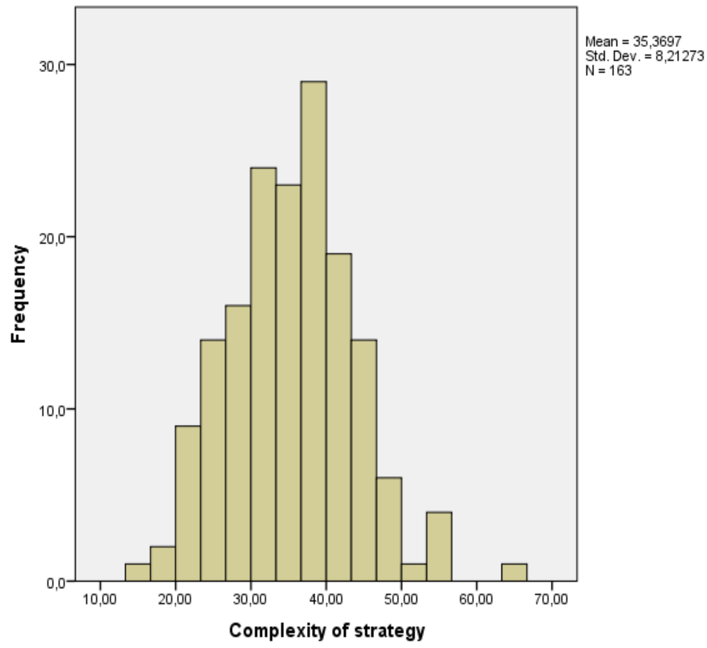
Table 11: Priority description terms with corresponding coded subfields

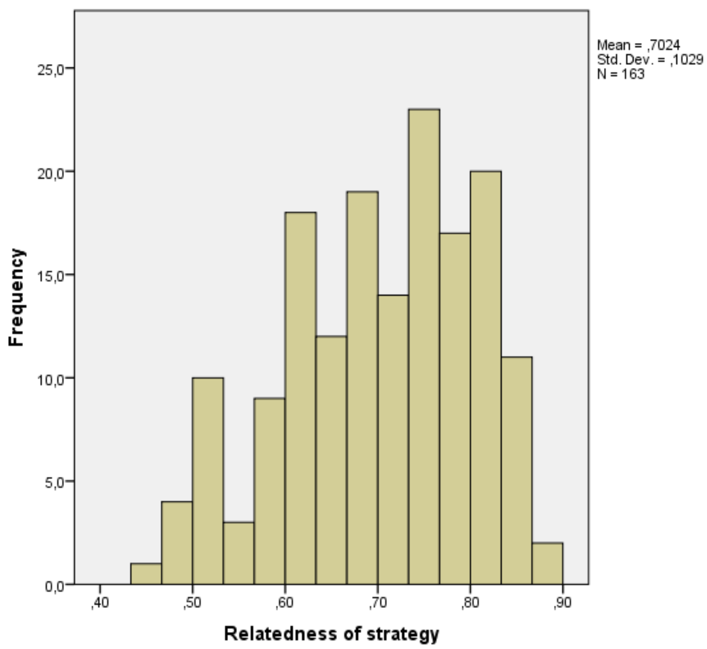
Priority term	Coded fields
Sustainability	Green & Sustainable Science & Technology
Transportation	Transportation Science & Technology
Sector cooperation	Industrial Relations & Labor
Construction	Construction + Civil engineering/Architecture
Food	Food science + Manufacturing/Agriculture
Innovation	Social sciences
Tourism	Hospitality, Leisure, Sport & Tourism
Waste management	Regional & Urban Planning + Environmental sciences
Waste	Environmental sciences
Logistics	Operations Research & Management Science + Transportation technology
Biomass	Biophysics + Energy
Green buildings	Green & Sustainable Science & Technology + Architecture/Construction
Energy (storage) systems	Energy + Automation & Control Systems
Bio-economy	Environmental sciences + Economics
Experience based industries	Theater/ Film, Radio, Television
Materials	Materials multidisciplinary
Creative industry	Art/Film, Radio, Television/Cultural science
Smart urban growth	Urban planning + Urban studies

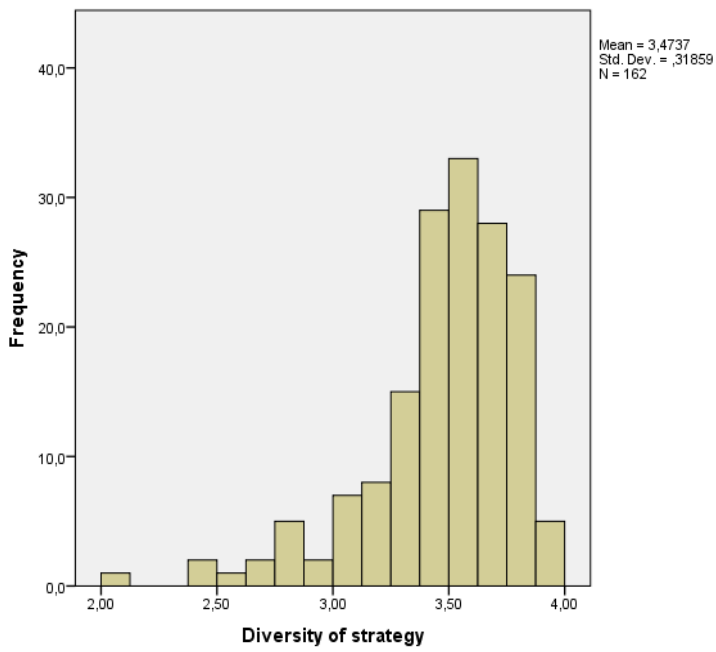
Appendix C: Data descriptives

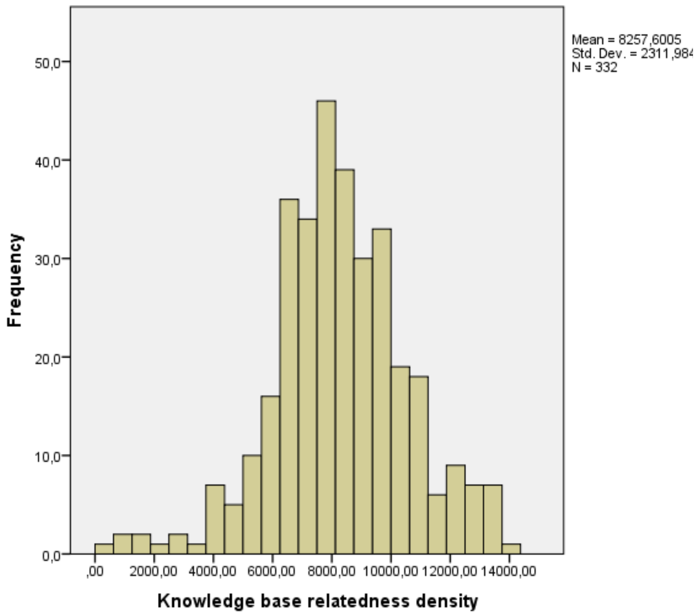
This appendix contains the histogram distributions for the variables used in this study.

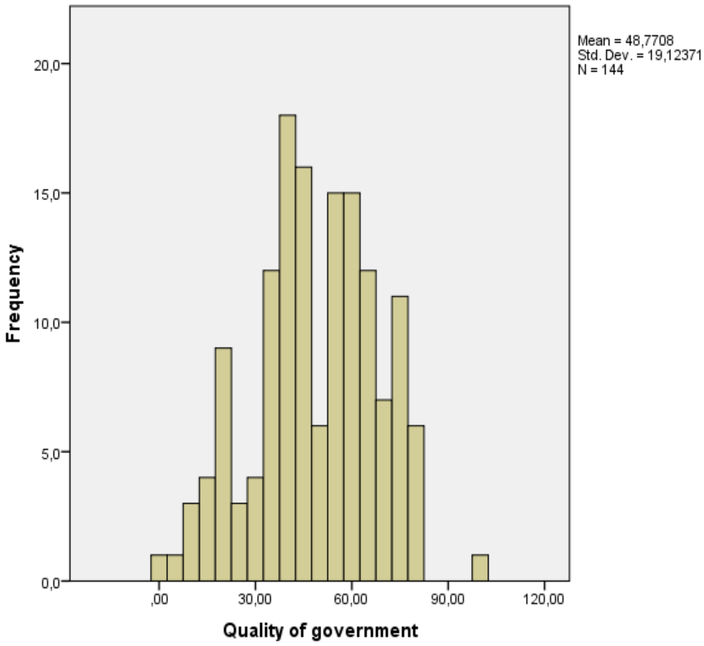


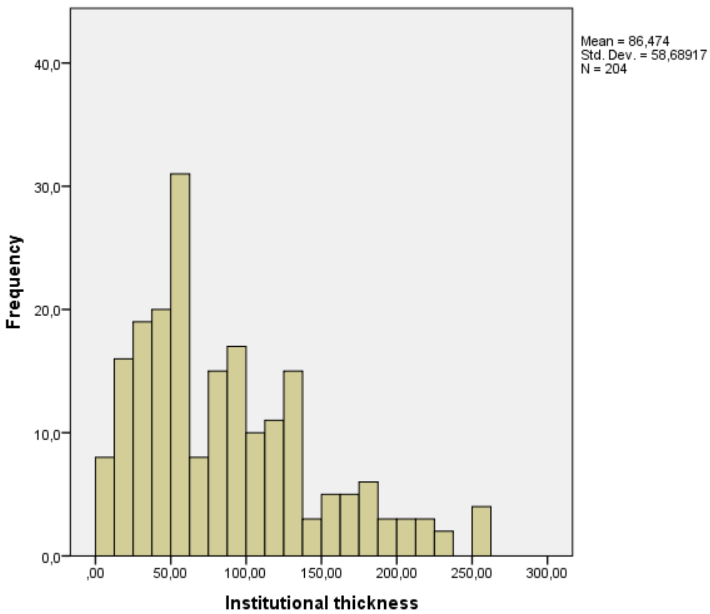


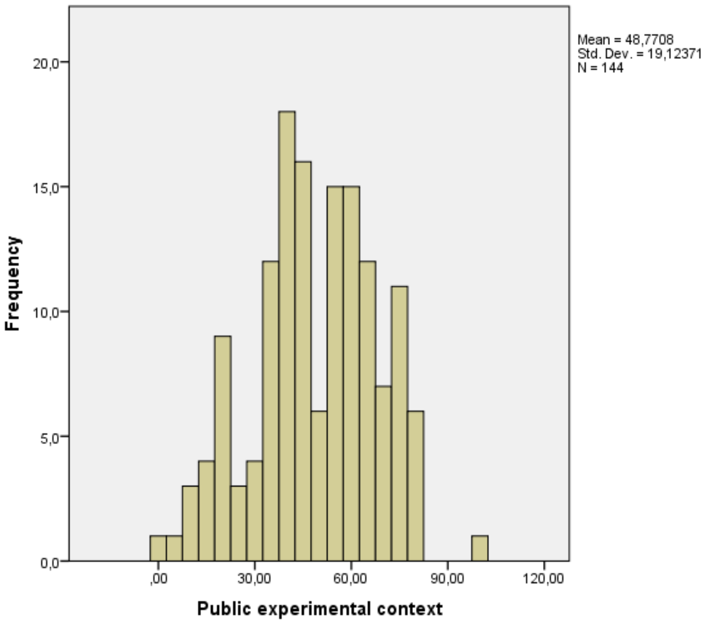


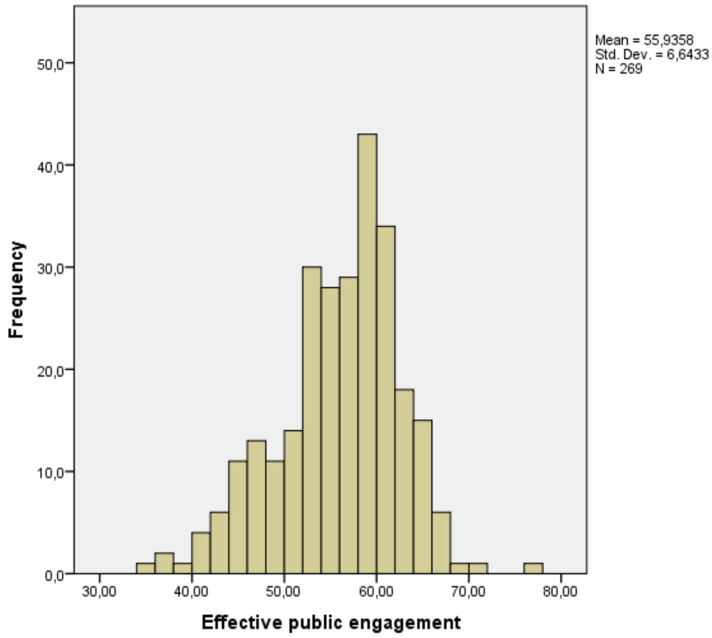


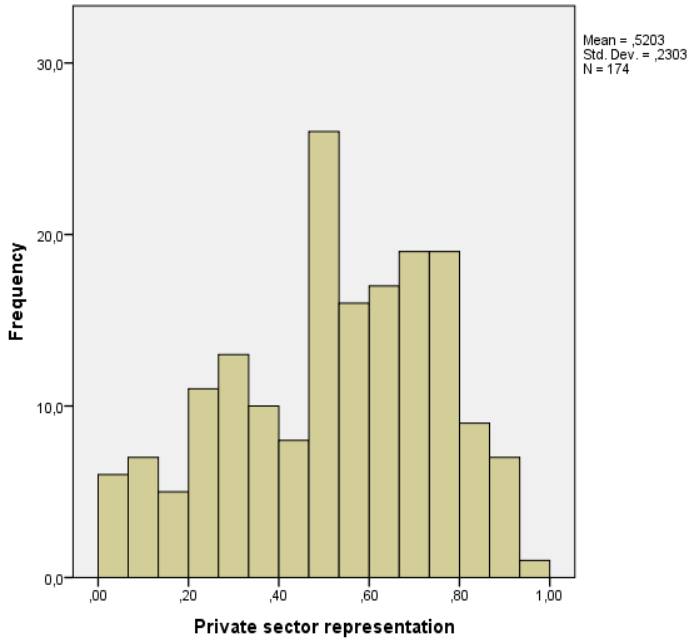

Mean = 8257,6005
Std. Dev. = 2311,984
N = 332
Frequency
50,0
40,0
30,0
20,0
10,0
0,0
,00
2000,00
4000,00
6000,00
8000,00
10000,00
12000,00
14000,00
Knowledge base relatedness density

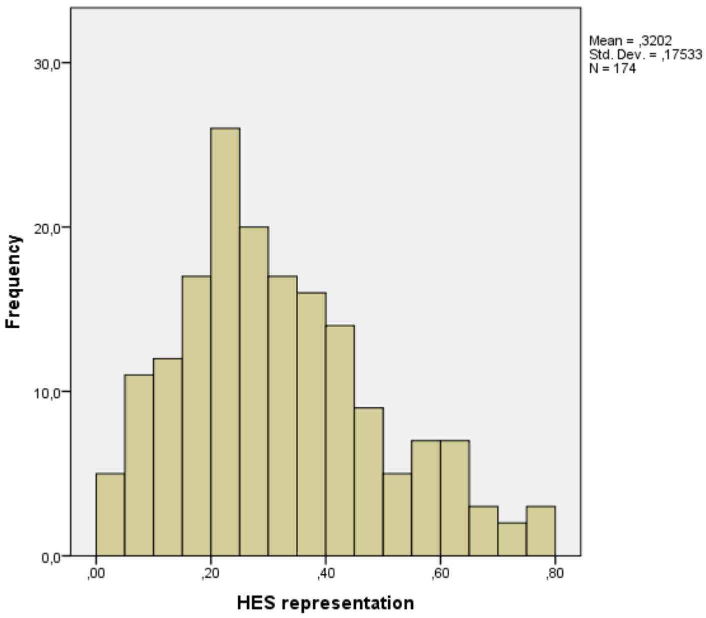

Mean = 48,7708
Std. Dev. = 19,12371
N = 144
Frequency
20,0
15,0
10,0
5,0
0,0
,00
30,00
60,00
90,00
120,00
Quality of government

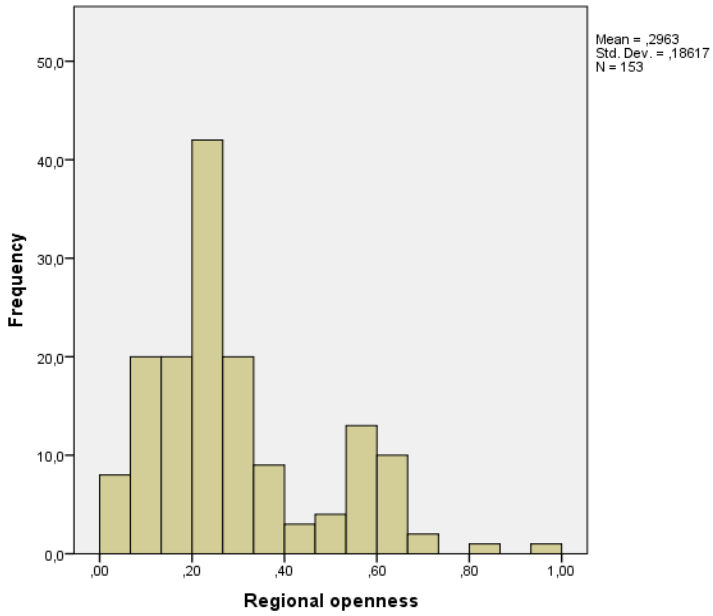

Mean = 86,474
Std. Dev. = 58,68917
N = 204
Frequency
40,0
30,0
20,0
10,0
0,0
,00
50,00
100,00
150,00
200,00
250,00
300,00
Institutional thickness

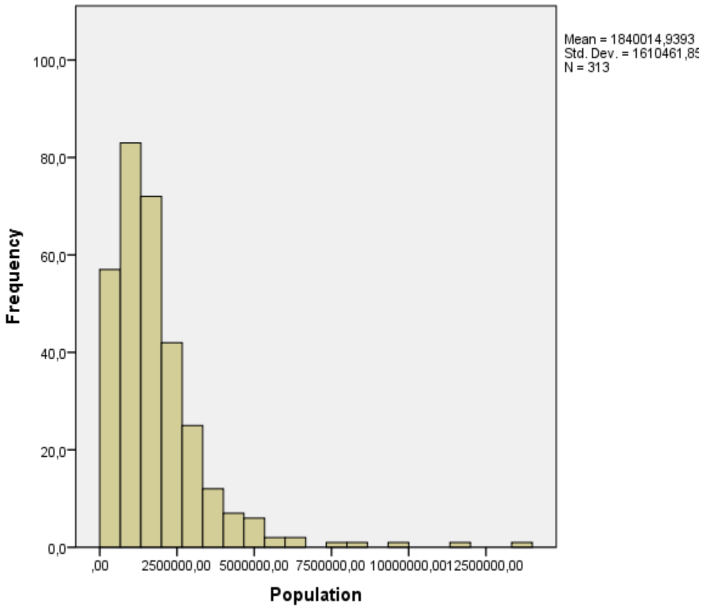

Mean = 48,7708
Std. Dev. = 19,12371
N = 144
Frequency
20,0
15,0
10,0
5,0
0,0
,00
30,00
60,00
90,00
120,00
Public experimental context

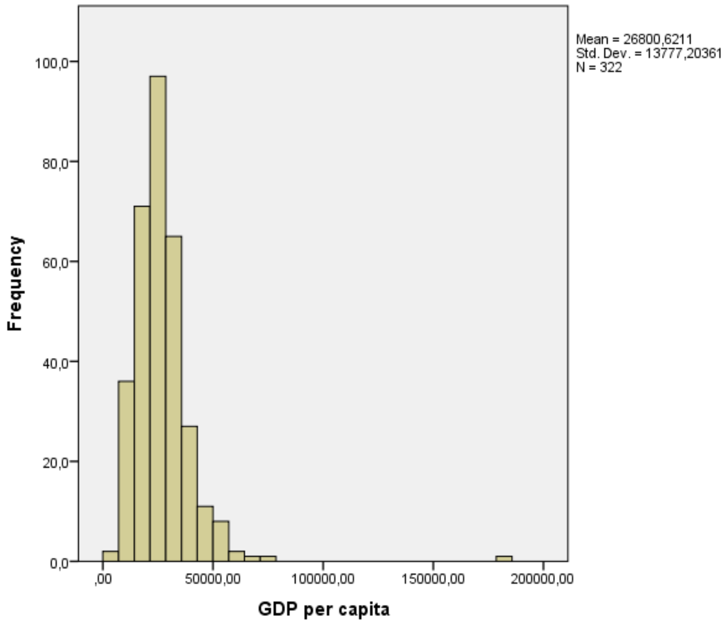

Mean = 55,9358
Std. Dev. = 6,6433
N = 269
Frequency
Effective public engagement

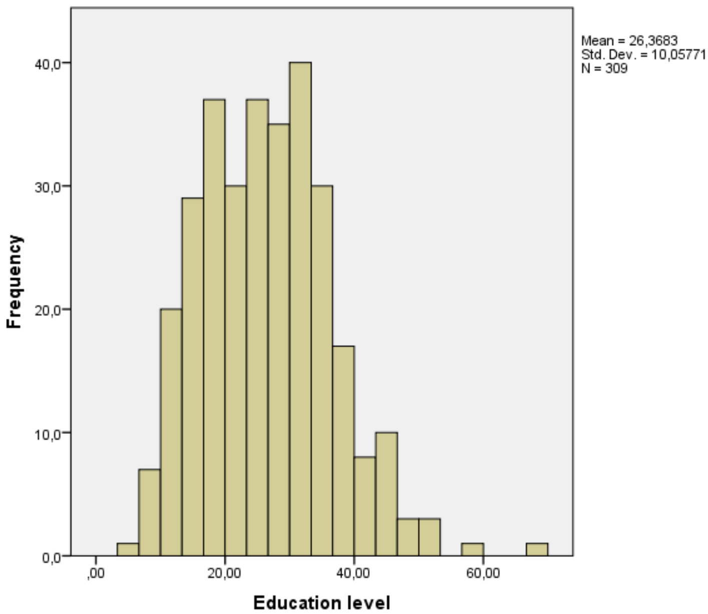

Mean = ,5203
Std. Dev. = ,2303
N = 174
Frequency
Private sector representation

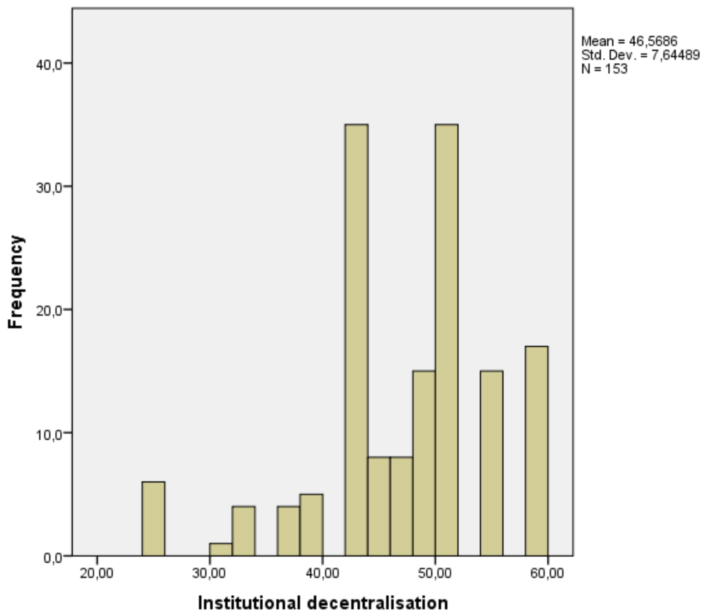

Mean = ,3202
Std. Dev. = ,17533
N = 174
Frequency
HES representation

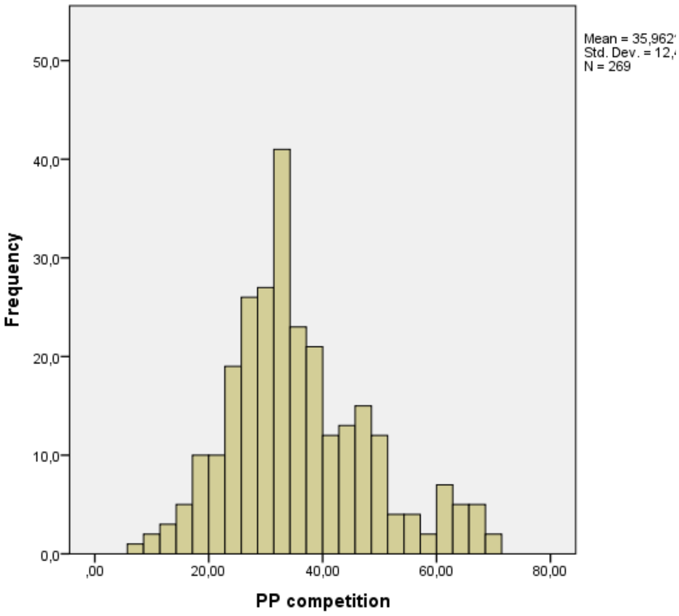

Mean = ,2963
Std. Dev. = ,18617
N = 153
Frequency
Regional openness

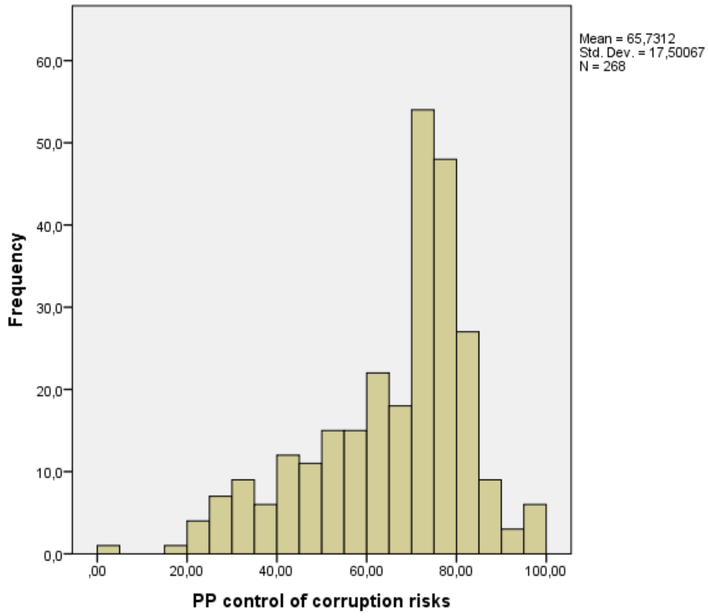

Mean = 1840014,9393
Std. Dev. = 1610461,85
N = 313
Frequency
Population

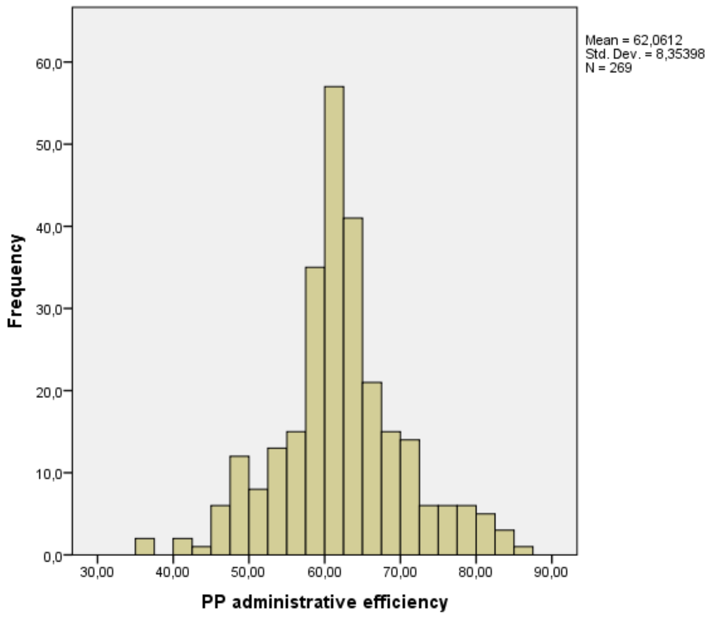

Mean = 26800,6211
Std. Dev. = 13777,20361
N = 322
Frequency
GDP per capita


Mean = 26,3683
Std. Dev. = 10,05771
N = 309
Frequency
Education level


Mean = 46,5686
Std. Dev. = 7,64489
N = 153
Frequency
Institutional decentralisation


Mean = 35,962
Std. Dev. = 12,
N = 269
Frequency
PP competition


Mean = 65,7312
Std. Dev. = 17,50067
N = 268
Frequency
PP control of corruption risks


Mean = 62,0612
Std. Dev. = 8,35398
N = 269
Frequency
PP administrative efficiency

-

Appendix D: Descriptive insights

In Figure 16 the non-normalised priority relatedness scores are depicted. A score of zero indicates that the relatedness score of the prioritisation strategy is lower than the overall relatedness density score of the regional knowledge base. A score of one indicates that the relatedness score of the prioritisation strategy is higher than the average relatedness density score of the regional knowledge base. Therefore, this measure gives an indication if regions prioritise above or below their average level of diversification opportunity.

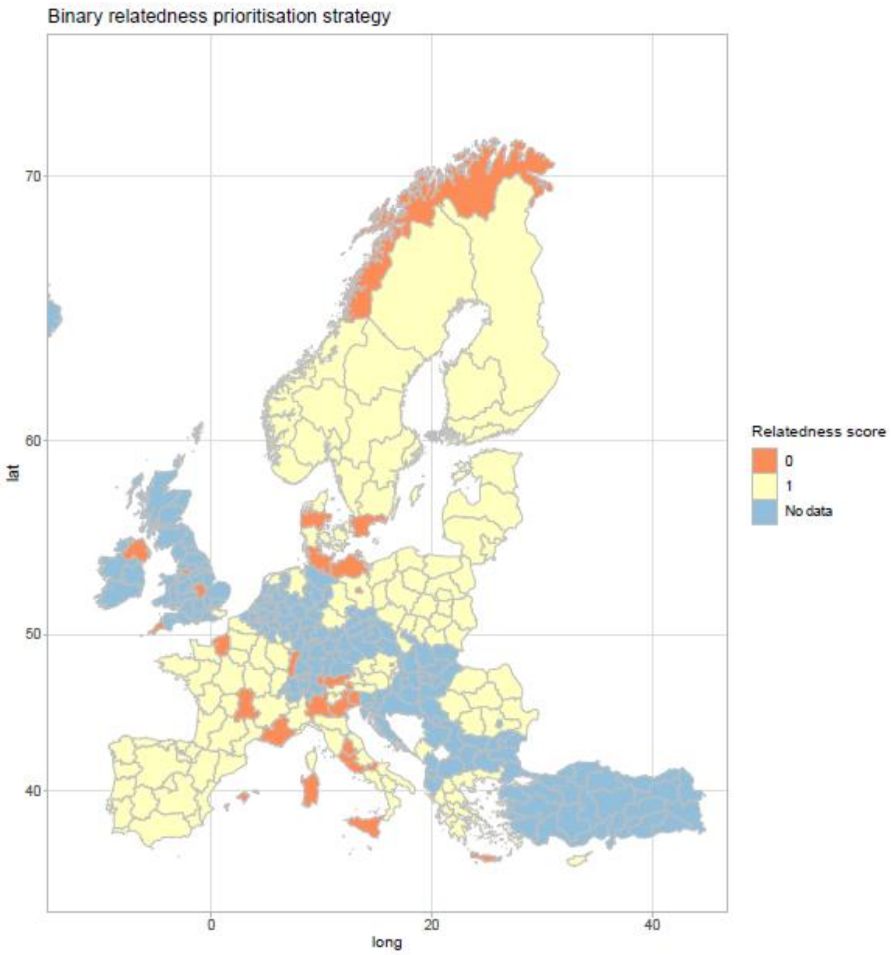


Figure 16: Map depicting if a region prioritised above (1) or below (0) its average level of relatedness density

<u>Priority spaces</u>

In Figure 17 the regional relatedness space is depicted. This figure illustrates the relative similarities of prioritised subfields between individual European regions. The closer two regions are placed together in this space, the similar the prioritisation strategy between the two regions. The regions are labelled with their respective NUTS 2 code.

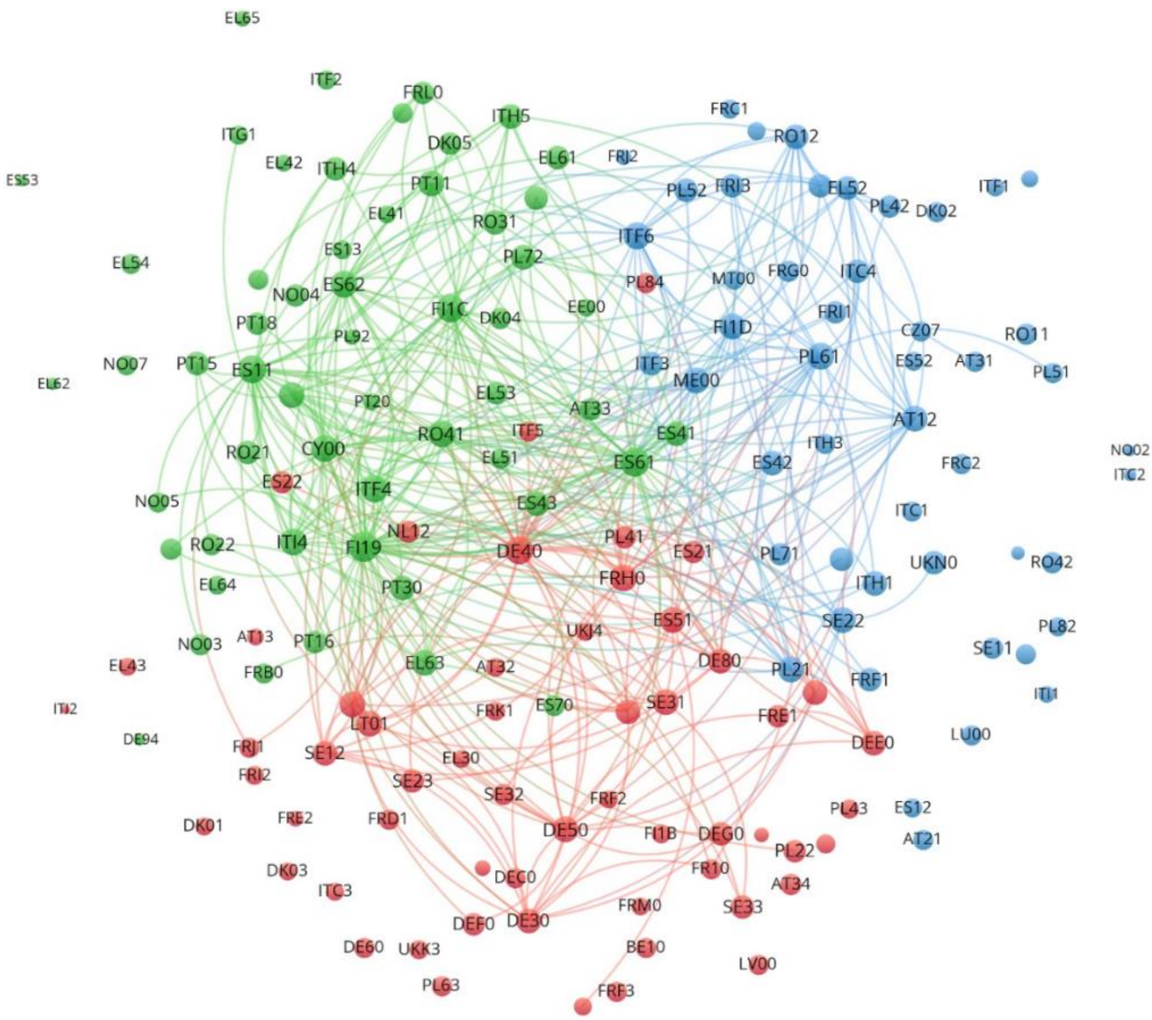


Figure 17: Regional relatedness space based on the prioritised subfields in the RIS3 program

Appendix E: Imputation statistics

This appendix contains statistics from the imputation procedure. Note, that in this procedure only the regions of which the priorities were collected are included. Figure 18 and Table 12 show further information on the missing data patterns in addition to Figure 1.

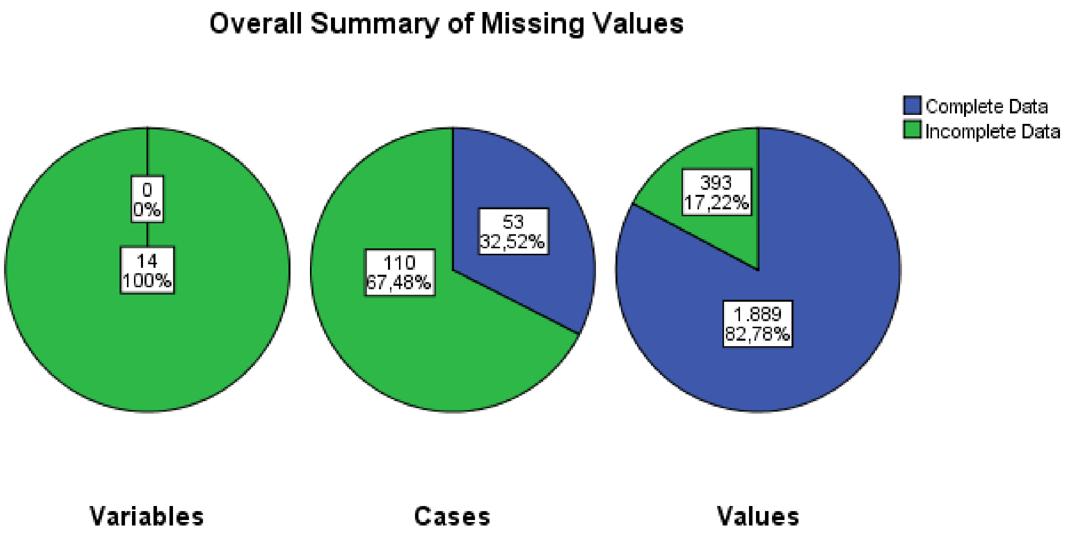


Figure 18: Overall summary of missing values of institutional variables for the regions subjected to the imputation procedure.

Table 12: Variable summary of the institutional variables for the regions subjected to the imputation procedure.

Variable Summary

	Missing		Valid N	Mean	Std. Deviation
	N	Percent			
Quality of governance	56	34,4%	107	49,4860	17,34737
HES representation	52	31,9%	111	,3577	,17256
Private sector representation	52	31,9%	111	,4648	,23646
Institutional thickness	45	27,6%	118	76,8373	54,45611
Institutional decentralisation	44	27,0%	119	47,5882	6,62739
Regional opennes	44	27,0%	119	,2261	,12657
Public experimental context	43	26,4%	120	,0019289079	,00244773122

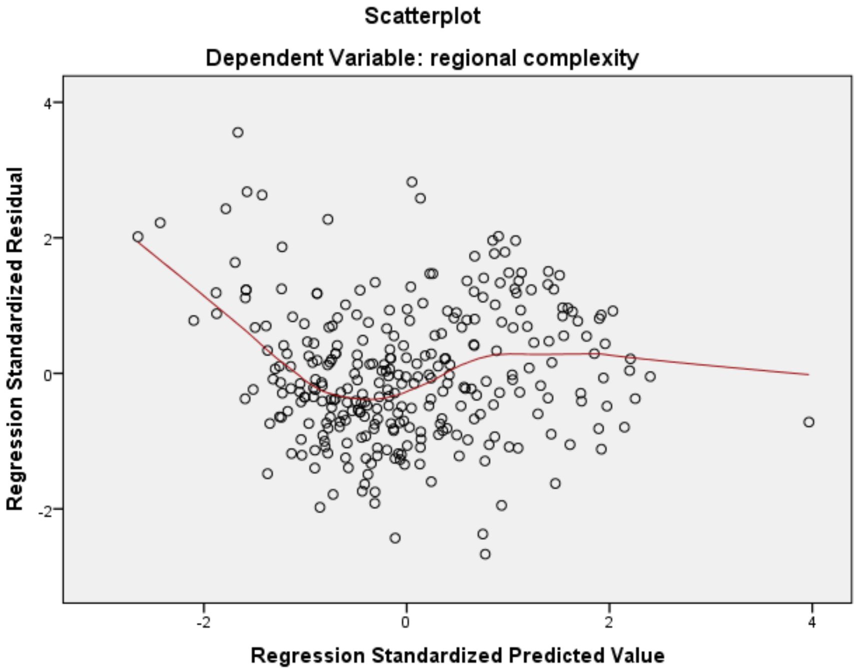
Appendix F: Regression statistics

In this appendix the quality indicators for the regression models are displayed. The residual plots for each model are presented, which were checked on homoscedasticity. Secondly, a multicollinearity indicator in the form of VIF-scores is given of which cases above two were investigated for problematic multicollinearity issues.

Model 1

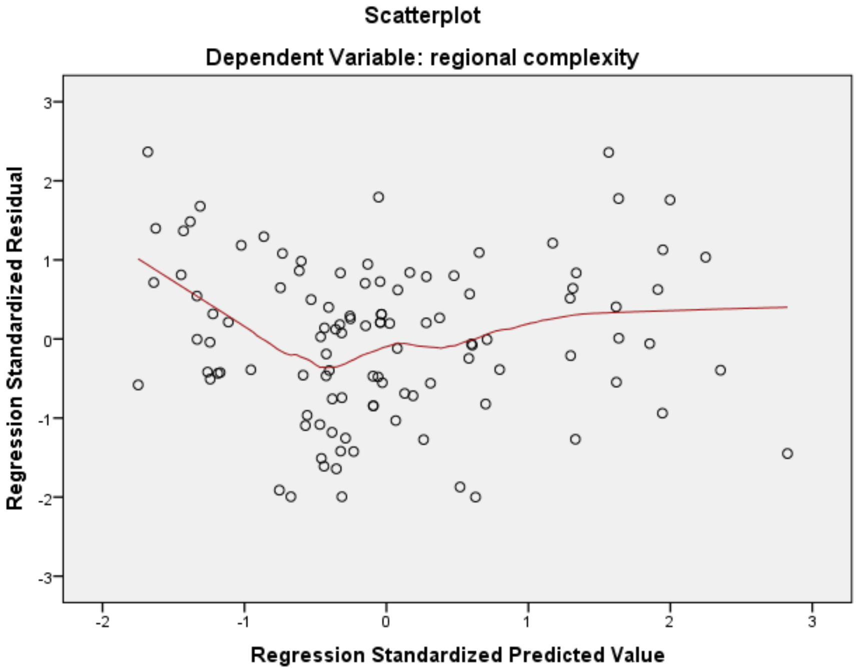
Knowledge base model

<u>Residuals</u>



Institutional model

<u>Residuals</u>



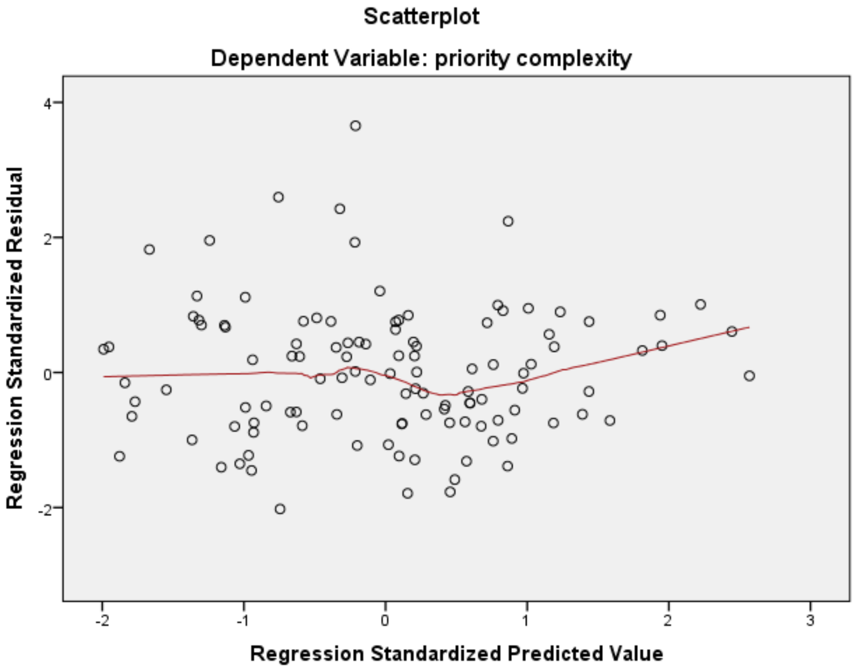
<u>VIF-scores</u>

Multicollinearity test model 1		
	Knowledge base	Institutional
Relatedness density	1,383	
Institutional thickness		2,499
Quality of Governance		1,869
Population	1,044	1,163
Gdp per capita	1,677	2,768
Education level	1,990	1,853

Model 2

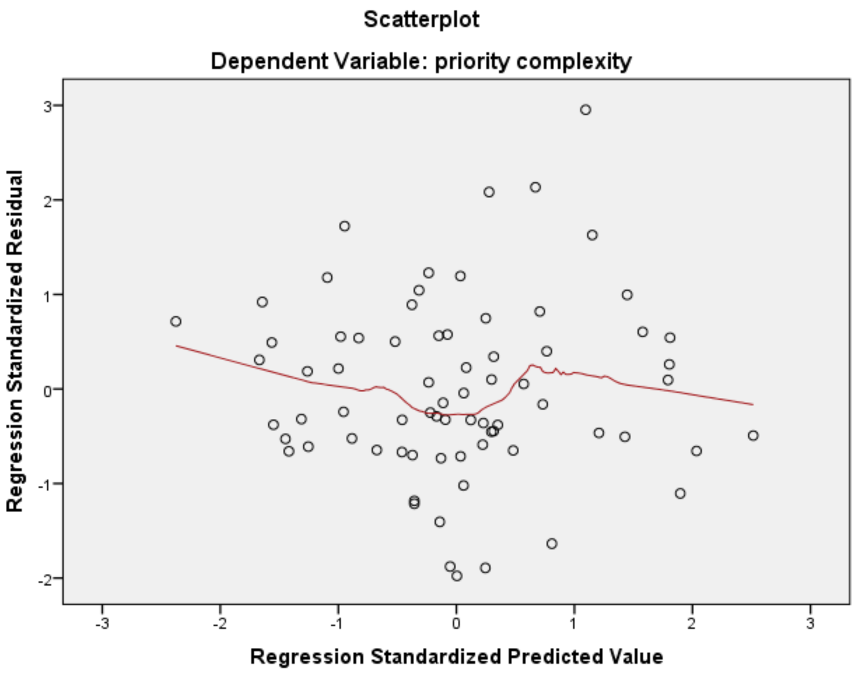
Knowledge base model

<u>Residuals</u>



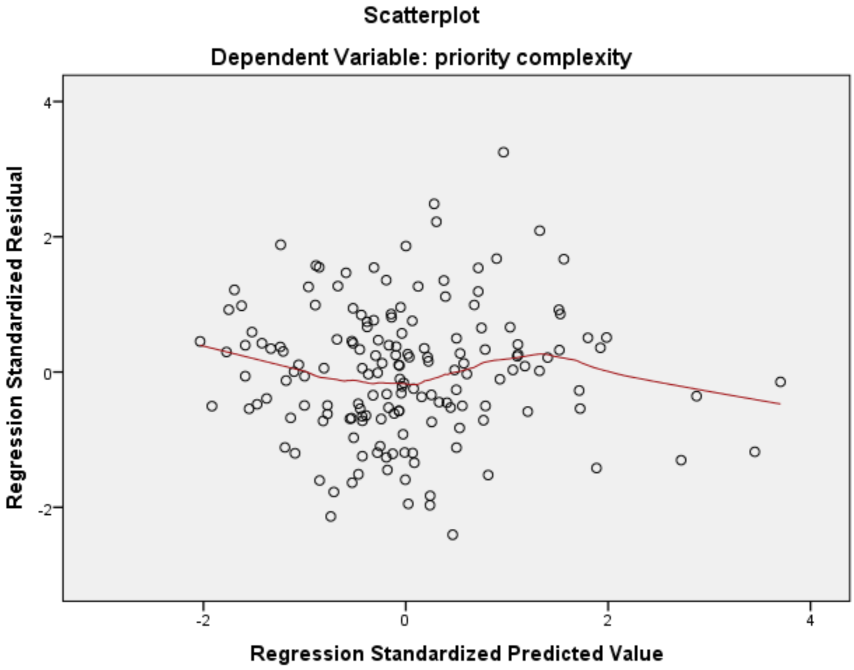
Institutional control model

<u>Residuals</u>



Institutional full model

<u>Residuals</u>



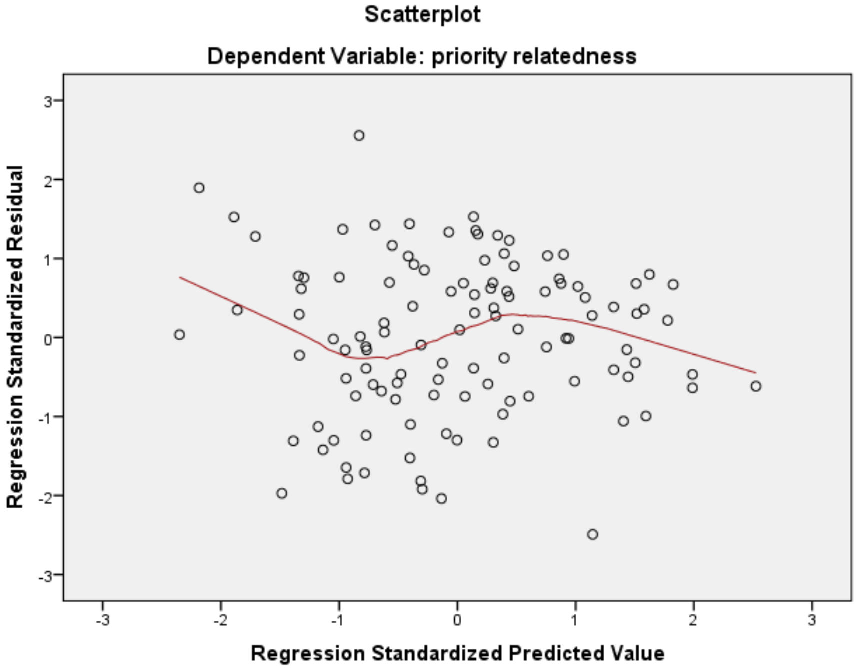
<u>VIF scores</u>

Multicollinearity test model 2.1			
	VIF score Knowledge base	VIF score Institutional (control)	VIF score Institutional (full)
Relatedness density	1,447		
Public experimentation		1,772	1,487
PP administrative efficiency		1,761	1,552
HES representation		1,281	1,180
Open exports		1,298	1,092
Population	1,010	1,309	1,426
Gdp per capita	1,301	2,342	1,580
Education level	1,571	1,916	1,895
Institutional decentralization	1,014	1,266	1,143

Model 3

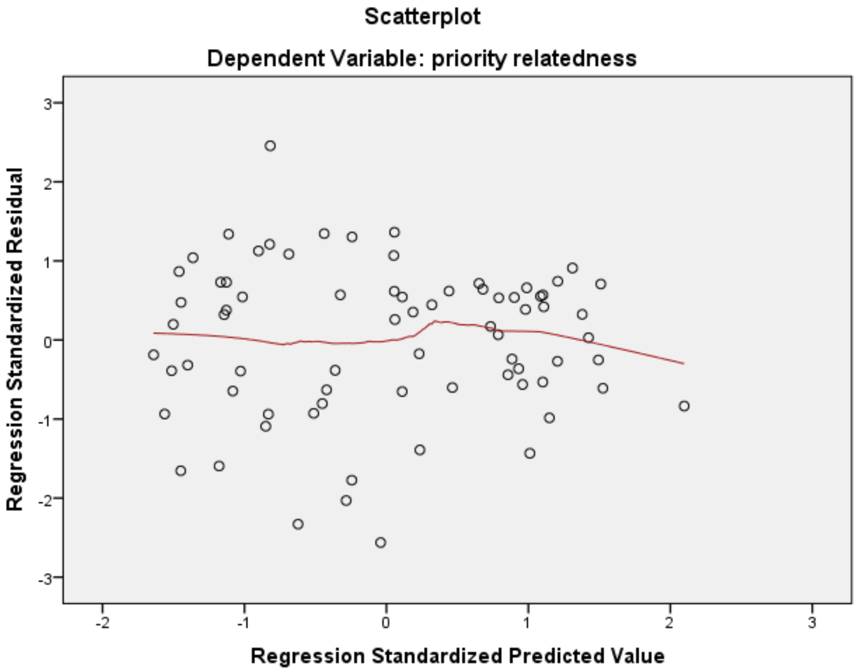
Knowledge base model

<u>Residuals</u>



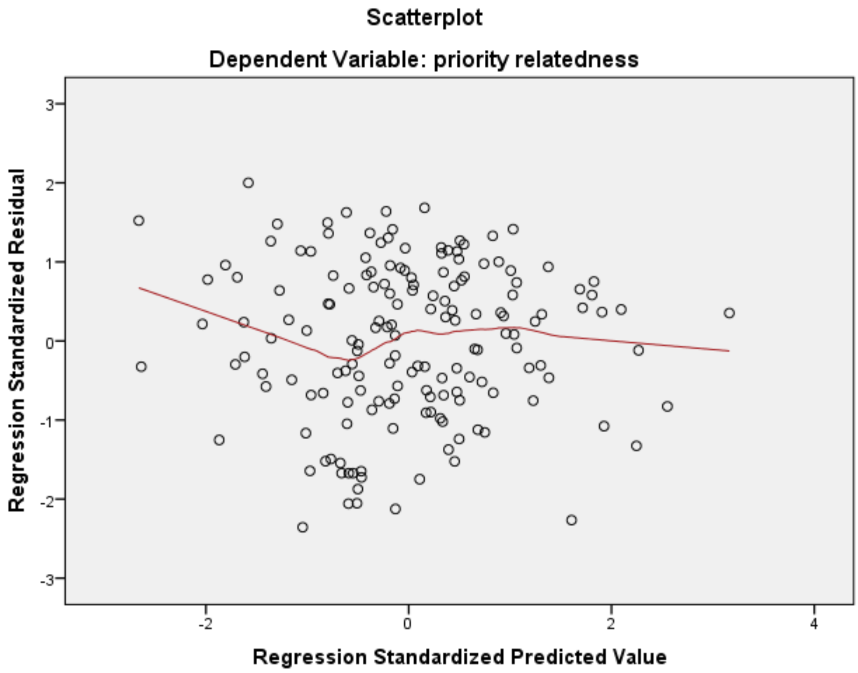
Institutional control model

<u>Residuals</u>



Institutional full model

Residuals



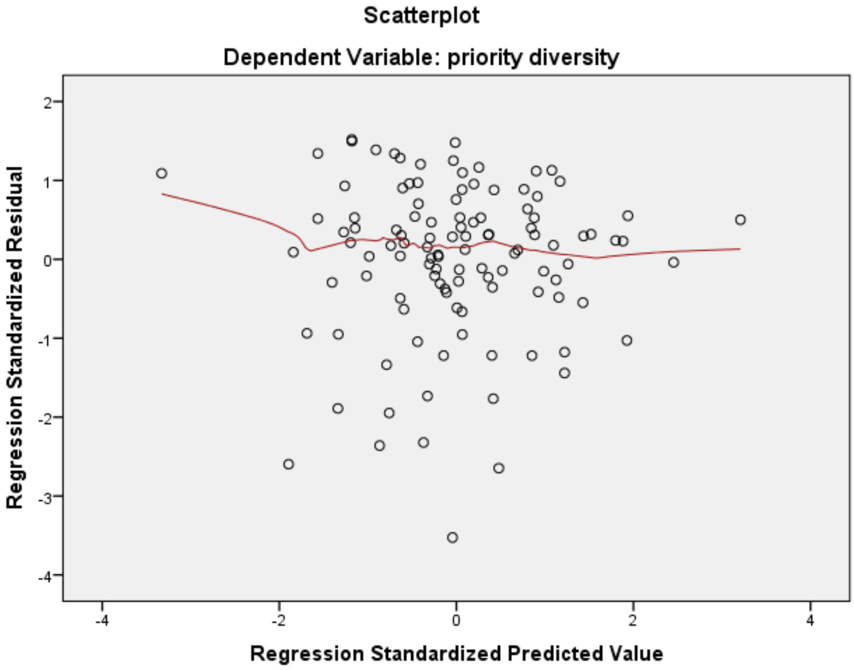
VIF scores

Multicollinearity test model 3.1			
	VIF score Knowledge base	VIF score Institutional (control)	VIF score Institutional (full)
Relatedness density	1,447		
PP governance		1,673	1,627
Business representation		1,236	1,261
Population	1,010	1,074	1,018
Gdp per capita	1,301	1,548	1,555
Education level	1,571	1,938	1,823
Institutional decentralization	1,014	1,096	1,084

Model 4

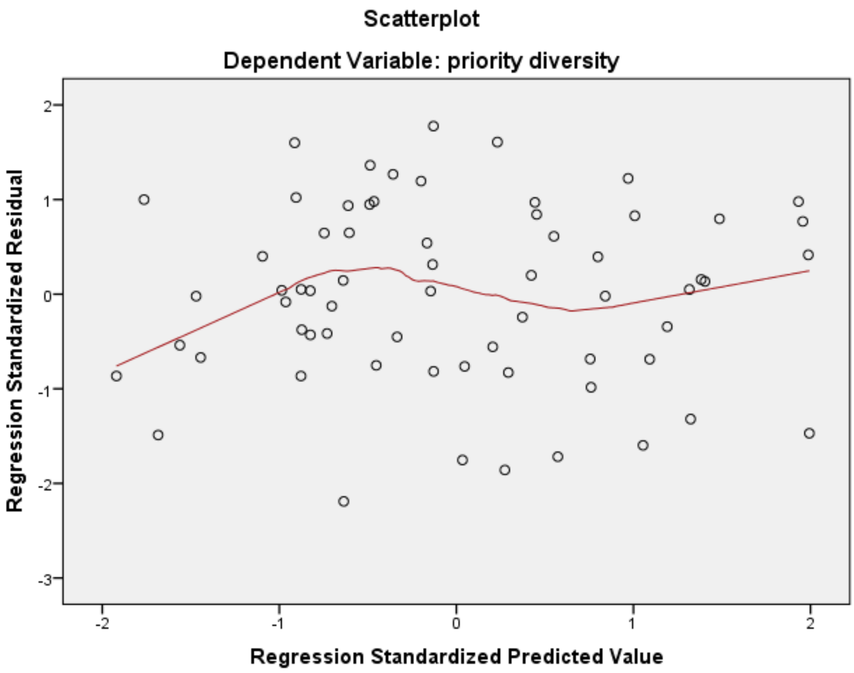
Knowledge base model

<u>Residuals</u>



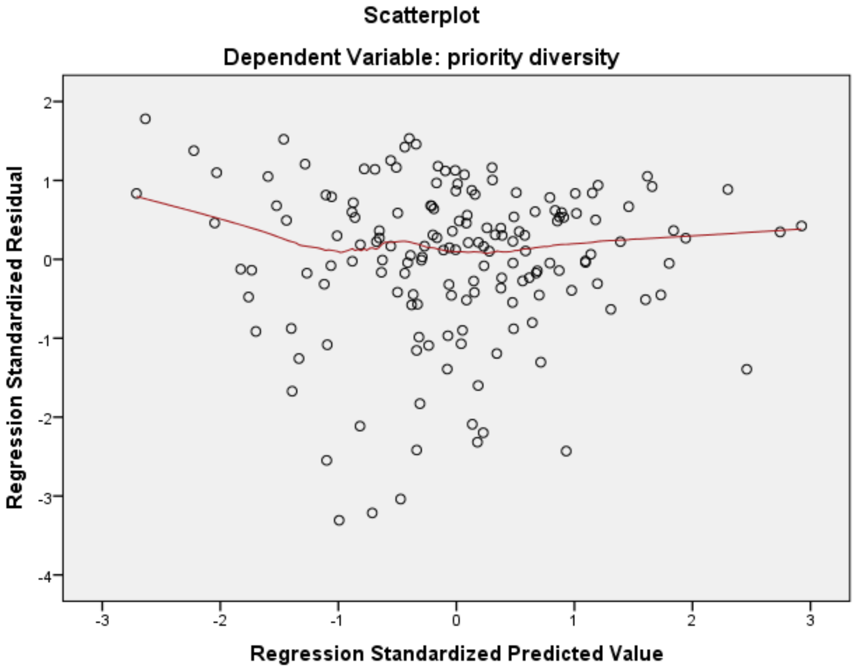
Institutional control model

<u>Residuals</u>



Institutional full model

Residuals



VIF scores

Multicollinearity test model 4.1			
	VIF score Knowledge base	VIF score Institutional (control)	VIF score Institutional (full)
Knowledge base complexity	1,437		
Public experimentation		1,706	1,448
HES representation		1,449	1,194
Open exports		1,876	1,083
Population	1,011	1,333	1,399
Gdp per capita	1,290	2,845	1,565
Education level	1,582	1,636	1,396
Institutional decentralization	1,011	1,166	1,109

Appendix G: Cluster statistics

In this appendix the ANOVA tables and iteration history of both cluster analyses is displayed. The ANOVA tables show that the values of the variables are significantly different across the clusters in both cluster analyses. The iteration tables show that in both cluster analyses significant cluster centres could be found within ten iterations.

Knowledge-institutional configuration cluster

ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Regional complexity	32,798	2	,603	160	54,435	,000
Quality of governance	56,963	2	,300	160	189,579	,000
Institutional thickness	55,133	2	,323	160	170,515	,000

Iteration History[a]

Iteration	Change in Cluster Centers		
	1	2	3
1	1,486	1,325	1,833
2	,229	,375	,315
3	,193	,000	,141
4	,176	,060	,155
5	,135	,000	,151
6	,103	,126	,067
7	,074	,125	,027
8	,049	,000	,057
9	,041	,000	,053
10	,000	,000	,000

Knowledge-institutional configuration with institutional prioritisation features cluster

ANOVA

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Quality of governance	48,295	2	,409	160	118,133	,000
Institutional thickness	39,740	2	,516	160	77,053	,000
Regional complexity	38,888	2	,526	160	73,876	,000
Effective public engagement	48,828	2	,402	160	121,420	,000
PP competition	18,120	2	,786	160	23,053	,000

PP control of corruption risks	49,422	2	,395	160	125,204	,000
PP administrative efficiency	35,885	2	,564	160	63,634	,000
Private sector representation	14,288	2	,834	160	17,134	,000
HES representation	6,833	2	,927	160	7,370	,001
Public experimental context	4,593	2	,955	160	4,809	,009
Regional openness	7,671	2	,917	160	8,369	,000

Iteration History[a]

Iteration	Change in Cluster Centers		
	1	2	3
1	4,076	3,910	4,515
2	,307	1,021	1,171
3	,171	,984	,606
4	,208	,587	,386
5	,211	,339	,257
6	,171	,306	,075
7	,247	,360	,054
8	,232	,212	,102
9	,092	,117	,000
10	,000	,000	,000