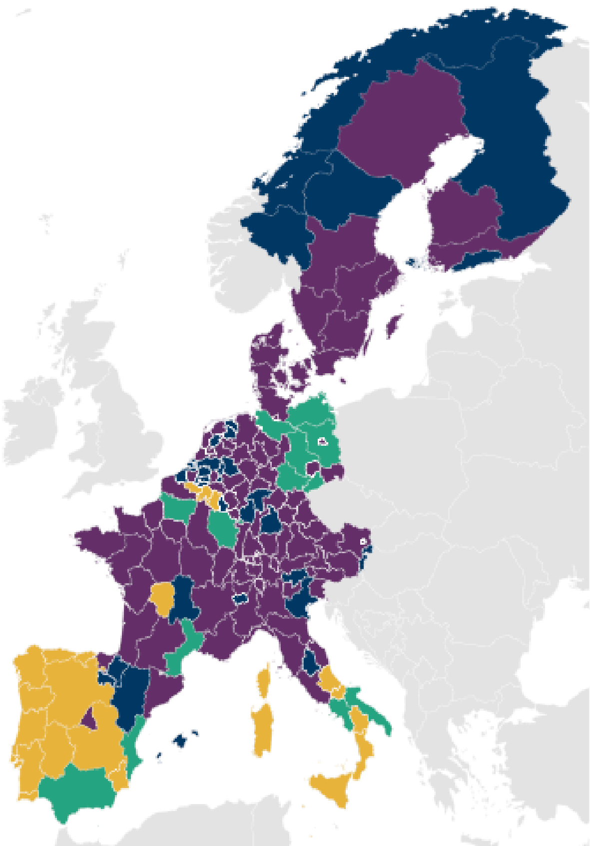
Technological Diversification in Peripheral Regions:

An Empirical Study on the Importance of Interregional Linkages across European Regions

How do the interregional linkages with core and peripheral regions affect the probability of a region to diversify?

Sebastiaan A. M. Kragting

Utrecht University



Universiteit Utrecht

dialogic
innovatie • interactie

Technological Diversification in Peripheral Regions:

An Empirical Study on the Importance of Interregional Linkages across European Regions

by

Sebastiaan A. M. Kragting

Contact	Student Number
S.A.M.Kragting@students.uu.nl	6935559

Supervisor I:	Dr. I. Wanzenböck
Supervisor II:	Dr. P. den Hertog
Examiner:	Dr. G. Heimeriks
Project Duration:	October, 2022 - June, 2023 (32 Weeks)
Field of Science:	MSc. Innovation Sciences
Faculty:	Geosciences, Dept. of Sustainable Development

Cover:	Entry of Technological Fields 1998-2018
LaTeX:	https://dzwaneveld.github.io
License:	https://creativecommons.org/licenses/by-nc/4.0/
GitHub:	https://github.com/SebastiaanK97/EconomicGeography

Universiteit Utrecht

dialogic innovatie • interactie

Preface

Technological Diversification in the Periphery:
An Empirical Study of the Importance of
Interregional Linkages across European Regions

The presented document is my thesis on technological diversification in the periphery which investigates how peripheral regions are able to diversify. The thesis has been written as closing research to graduate from the masters programme Innovation Sciences. The engagement period of the research lasted from October up to June 2022.

The research is undertaken under supervision of Utrecht University whereas the topic has emerged out of my own interest in technological development in peripheral regions. My specific interest lies in how technology as a social concept impacts society and therefore must be treat sustainable and inclusive in our economies. The research has been intriguing, and during the period I've learned plenty on how technology and economy shape together.

Therefore, I would like to address special thanks to my supervisors for their immense support during the process. I both thank Dr. I. Wanzenböck and Dr. P. den Hertog for their in-depth discussions and feedback on my process. You have raised the level of my thesis by a lot and I was eager to learn from your thoughts.

I also thank my peer students from the masters programme for the time we went through and the results we were able to deliver together. I appreciate the time spent together during the courses and of course I hope to keep seeing you as a friend and during our career. I would also like to thank my family for their stable support and caring when in need. The last year was challenging but marks an astonishing ending to my studies.

I hope the thesis is entertaining and engrossing to read.

Sebastiaan A. M. Kragting
Utrecht, June 2023

Summary

Technological Diversification in the Periphery: An Empirical Study of the Importance of Interregional Linkages across European Regions

This thesis puts the literature on Evolutionary Economic Geography central and assumes that technological diversification is necessary in peripheral regions to foster sustainable and inclusive growth across regions. However, the paradox is that peripheral regions are less probable to diversify because they tend to lock-into their existing specialisations. Therewith, the knowledge production in peripheral regions is low, and due to the principle of relatedness, peripheral regions are prone to lock-into their narrow knowledge base. The literature on economic geography suggests that interregional linkages leads to external knowledge spillovers and therefore complements internal knowledge production. Thus, interregional collaboration is especially important for peripheral regions to escape a lock-in into their existing specialisations. This thesis scrutinises this general perception on interregional linkages by explaining that core and peripheral regions receive different benefits from external knowledge spillovers. The main research question is: *How do the interregional linkages with core and peripheral regions affect the probability of a region to technologically diversify?*

To answer this question the thesis adheres to a quantitative explanatory and exploratory approach to understand technological diversification in core and peripheral regions. The analyses draws upon economic and patent data between $2005 - 2018$ to investigate the entry of technological fields. The conclusions are as following. First, regression models suggest that external knowledge spillovers from core regions are of greater importance to technological diversification than that from peripheral regions. Further analysis of assortative mixing shows that core regions increasingly collaborate with core regions whilst peripheral regions remain rigid and do not increase in interregional collaboration. Therefore, the importance of interregional collaboration seems promising in theory but might not entirely enable peripheral regions to catch-up to core regions. Nevertheless, external knowledge spillovers through collaboration remain crucial for technological diversification in peripheral regions. Second, a network analysis of co-invention reveals that those peripheral regions that do diversify are exceptionally high in betweenness centrality. Likely, these regions can function as a bridge between core and peripheral regions, for instance, due to providing access to different labour markets (and migration), or flows of natural resources. Third, whilst peripheral regions do show low novelty in technological diversification, an unexpected novel finding is that core regions that do collaborate with peripheral regions show highest novelty in technological diversification. Therefore, potentially, the collaboration between core and peripheral regions induces recombination of knowledge that is new to the world.

Contents

Nomenclature

The following nomenclature describes the abbreviations and mathematical symbols used.

Abbreviations

Abbreviation	Definition
EU	European Union
EEG	Evolutionary Economic Geography
GDP	Gross Domestic Product
S3	Smart Specialisation Strategy
SNA	Social Network Analysis

Symbols

Symbol	Definition	Unit
$AC_{r,t}$ ...	Absorptive Capacity	[1]
$ILC_{r,t}$ ...	Interregional Linkages with Core Regions	[1]
$ILP_{r,t}$ ...	Interregional Linkages with Peripheral Regions	[1]
$RD_{i,r,t}$ ...	Relatedness Density	[1]
$RTA_{i,r,t}$ ...	Relative Technological Advantage	[1]
$\varphi_{i,j,t}$ ...	Relatedness	[1]
ϕ_c ...	Country Fixed-Effects	[1]
ψi ...	Technological Class Fixed-Effects	[1]

1
Introduction

”In the European Union (EU) in the new millennium, inequality among regions has turned sharply up” (Iammarino et al., 2019, p. 2). The income inequality across regions defines the core-periphery dichotomy in the EU[1], and this regional inequality continues to increase. Herewith, peripheral regions fall behind in economic development (GDP per capita) in opposition to core regions. Therewith, the peripheral regions decline in population, reduce in employment, and suffer from unemployment in opposition to core regions. Iammarino et al. (2019) call for evidence to promote sustainable and inclusive growth across regions.

The literature on evolutionary economic geography (EEG) describes technological diversification as a key driver towards economic growth (Dosi, 1982). Technological diversification refers to a region specialising into technological fields that are new to the region (Hassink & Gong, 2019). Therefore, technological diversification leads to new activities which foster employment and suppress unemployment growth (Castaldi et al., 2015; Frenken et al., 2007).

The production of knowledge is seen as central to technological diversification. The production of knowledge is a cumulative, path-dependent, and interactive process (Boschma et al., 2015). In other words, regions produce knowledge that is related to their existing knowledge base, and thus technological diversification is not a random process. Therefore, the production of knowledge concentrates in space and time; meaning that core regions easily exploit knowledge that is related to their diverse knowledge base whilst peripheral regions struggle to exploit knowledge from their narrow knowledge base (McCann & Ortega-Argilés, 2015). The process of technological diversification towards related knowledge defines the principle of relatedness. According to Hidalgo et al. (2018), ”the principle of relatedness is a force that increases spatial inequality and can reduce the ability of peripheral [regions] to develop” (p. 455).

The paradox is that peripheral regions require to diversify for sustainable and inclusive growth, however, peripheral regions are less probable to diversify than core regions, and more likely to lock-in into existing specialisations, due to the principle of relatedness (Hassink & Gong, 2019; Iacobucci & Guzzini, 2016).

The literature on economic geography suggests that regions draw upon external knowledge to complement internal knowledge production (Bathelt et al., 2004). Several observations have been made that interregional linkages i.e., collaboration between regions by organisations, causes external knowledge to spillover between regions and reduces the tendency of regions to lock-in into their knowledge base (Boschma & Iammarino, 2009; Grillitsch & Nilsson, 2015). Accordingly, the observation by Balland and Boschma (2021) is that interregional linkages indeed enable (especially peripheral) regions to diversify.

This thesis builds further upon the importance of interregional linkages by Balland and Boschma (2021) in the body of literature on EEG under the header of smart specialisation by differentiating core and peripheral regions further. Within smart specialisation, McCann and Ortega-Argilés (2015) discuss that peripheral regions are lagging regions which follow core regions where knowledge production is higher. Therefore, peripheral regions need to link to core regions to foster external knowledge spillovers to induce technological diversification (Balland & Boschma, 2021; Iacobucci & Guzzini, 2016). As McCann and Ortega-Argilés (2015) suggest, public policy must maximise 'learning linkages' between core and peripheral regions. This means to transfer knowledge from core to peripheral regions. Here, the implicit assumption is that interregional linkages with core and peripheral regions are of unequal importance (Hassink & Gong, 2019). Therefore, the main research question is:

How do the interregional linkages with core and peripheral regions affect the probability of a region to technologically diversify?

The research design is split into an explanatory and exploratory approach and draws upon regions in the EU for quantitative empirical analysis.

For the former, a regression model is used to estimate the effect of interregional linkage on the probability of a region to diversify. Here, a novelty is made by measuring interregional linkages with core and peripheral regions in Europe separately because the research seeks to explain whether external knowledge spillover yields unequal effects across the core-periphery dichotomy. This separation has been neglected in prior studies. We need to know this to confirm whether interregional linkages are equally important for core and peripheral regions to engage in technological diversification.

For the latter, the question is how interregional linkages establish in space and develop over time. Whether peripheral regions link mostly to peripheral regions, or, whether peripheral regions also link to core regions is unknown. In technical terms, is interregional linkage as a network homophilous or heterophilous (by assortative mixing) between core and peripheral regions. We need to know this to understand how core and peripheral regions can draw upon external knowledge in a time of increasing globalisation.

The research studies technological diversification in regions over $2005 - 2018$ over five subsequent periods. The research scope is delimited to ~ 176 regions by the Nomenclature of Territorial Units for Statistics (NUTS2) in Western European countries.[1] In line with the EU Cohesion policy, the peripheral regions fall below a 75% of the average GDP per capita, if else, the regions are considered as core regions. The research operationalises ~ 650 technological fields by the Cooperative Patent Classifications (CPC4). Therewith, the unit of analysis is at the region-technology level i.e., each technological field is observed solitary within each region.

[1]The delimitation of the sample is a result of combining economic with patent data. First, due to the availability of economic data in Eurostat, the sample consists of the 27 EU member states as of 2023, but also, Albania, Ireland, Lithuania, Montenegro, North Macedonia, Norway, Serbia, Swiss, and Turkey. Second, due to the eligibility of patent data in OECD, the sample is further delimited to Western European regions as explained in Appendix C.

The scientific relevance of the research holds in the argument to nuance the importance of interregional linkages, and therewith understand the bottlenecks in external knowledge spillovers between core and peripheral regions. In particular, the importance of interregional linkages is promising in theory, however, interregional linkages may pertain unequal over the core-periphery dichotomy empirically (Hassink & Gong, 2019). Further, an attempt is made to scrutinise how peripheral regions can be innovative through analysing peripheral regions that do diversify in their geography and as a network of co-invention (Eder, 2019).

The societal relevance of the research holds in examining technological diversification as a process of smart specialisation. Therewith, smart specialisation is a place-based (bottom-up) approach that sets technology-driven priorities to solve socio-economic challenges. In line with the European Investment Bank (EIB), the lesser developed (i.e., peripheral) European regions require to diversify in order to induce economic competitiveness, expand the labour market, and transition towards a sustainable and digital economy (Balland et al., 2019). In order to justify smart specialisation strategy (S3), peripheral regions should be considered as regions with valuable, but diverging, local capabilities (Balland & Boschma, 2021). If the process of technological diversification could be understood more comprehensively in terms of interregional linkages over core and peripheral regions, then, S3 is able to target technological change between core and peripheral regions more accurately, and therewith foster economic cohesion in the European Union. In summary, emphasising the bottlenecks between interregional linkage and regional diversification in peripheral regions allows justified allocation of funding for the EU Cohesion Policy (Foray, 2014).

The structure of the thesis is as following. Hereafter, Chapter 2 reviews the theoretical foundations of the core-periphery dichotomy, the principle of relatedness, and the importance of interregional linkages. Subsequently, Chapter 3 explains the methodology in terms of the data and the analytical approaches taken. Thereafter, Chapter 4 presents the empirical findings, which is followed by a conclusion on the key findings in Chapter 5. Finally, Chapter 6 describes the theoretical contributions, limitations, and policy implications of this thesis.

2
Theoretical Foundations

2.1. The Core-Periphery Dichotomy

This thesis defines a core-periphery dichotomy in order to contrast policy objectives over two types of regions i.e., core and peripheral regions (Iammarino et al., 2019). Therewith, the dichotomy is useful to separate further definitions of technological diversification, economic development, and external knowledge spillovers (Hassink & Gong, 2019).

The core-periphery dichotomy by McCann and Ortega-Argilés (2015) generalises how all the peripheral regions suffer from an inability to produce knowledge in a variety of dimensions (e.g., sectoral, transactional, technological, behavioural, financial, cultural) as opposed to core regions. The argument is that peripheral regions lack the required organisational conditions to induce sustainable and inclusive growth (Whittle & Kogler, 2020). This thesis builds upon the core-periphery dichotomy to generalise and explain how core and peripheral regions differ in their internal knowledge production. In particular, the core-periphery dichotomy summarises innovation in six stylised facts that dichotomise core and peripheral regions. The six stylised facts explain how organisations (e.g., firms, universities, and public institutions) exploit the knowledge that a region exhibits (Buenstorf & Klepper, 2009; Klepper, 2007).

The six stylised facts of innovation between the core-periphery dichotomy are as following. First, innovation tends to be higher in densely than in scarcely populated regions. Second, innovation tends to be higher in regions with diverse sectors than in regions that specialise. Third, innovation tends to be higher in regions with a variety of small firms rather than regions with few large firms. Fourth, innovation tends to be higher in regions with internationally oriented multinational firms. Fifth, innovation tends be higher in regions with high market potential. Sixth, the adoption of information and communication technologies emphasises the core-periphery dichotomy. As a consequence of these stylised facts, core regions thrive in knowledge production, whilst peripheral regions fall short (McCann & Ortega-Argilés, 2015). This thesis will build on, and scrutinise, the core-periphery dichotomy emphasised by these six stylised facts related to the process of technological knowledge production.

Last but not least, a novel stream in literature on the geography of innovation attempts to identify how peripheral regions provide benefits towards innovation (Eder, 2019). Peripheral regions may be innovative (Eder & Trippl, 2019) due to, for example, lower wages that are attractive for labour-intensive industries (and their corresponding technological fields), proximity to natural resources (due to less urbanisation) and industry specific advantages, or, peripheral regions limit unintentional knowledge spillover. Nevertheless, few generalisations have been made within this stream of literature, therefore, this thesis attempts to identify how peripheral regions are innovative in terms of technological diversification.

2.2. The Principle of Relatedness

The literature on EEG acknowledges that economic activities do not arise spontaneously, but instead arise as a path-dependent process of related activities (Dosi, 1982; Frenken et al., 2007). Therefore, EEG puts the principle of relatedness central to explain how regions diversify into activities that evolve out of their existing knowledge base. The theory suggests that technological diversification is not a random process. Instead, technological diversification in regions is dependent on the path they follow from the past, and thus, the specialisation into new technological fields diverges from region to region (Boschma et al., 2015; Rigby, 2015). In short, regions tend to diversify into technological fields that are alike to their hitherto present technological fields in their economy (Boschma et al., 2013).

In this thesis a technological field refers to an aggregation of single inventions among their knowledge producing organisations in an identical field and their corresponding knowledge base (Balland et al., 2019). Thus, diversification in a region refers to a new specialisation into a (to the region) novel knowledge base of a technological field (Hassink & Gong, 2019).

In this thesis, relatedness refers to how technological fields in a region co-occur. If technological fields co-occur frequently within regions, then technological fields are related between one another. Thus, relatedness refers to the implicit interconnections between technological fields as a consequence of co-occurrence. Therewith, the recombination of existing knowledge in technological fields should lead regions to diversify into technological fields that are implicitly related by co-occurrence (Boschma et al., 2015).

The principle of relatedness is further extended to the concept of the technology space. The technology space is the representation of relatedness between technological fields as a network. Therewith, related technological fields tie together, whereas unrelated technological fields do not tie and are distant in the technology space (Hidalgo et al., 2007). All in all, the argument is that regions diversify into technological fields that are nearby to their existing knowledge base, because, organisations tend to exploit nearby knowledge instead of knowledge that is distant to their existing knowledge base (Boschma et al., 2015).

The issue is that peripheral regions cannot draw from technological relatedness equally as core regions because peripheral regions tend to specialise narrowly (Caragliu et al., 2016). As Balland et al. (2019, p. 1263) state ”peripheral regions provide one of the most complicated cases to build an effective smart speciali[s]ation policy”. In particular, diversification for peripheral regions requires them to take ’long jumps’ from their existing knowledge base towards the novel knowledge in a new technological field. However, peripheral regions often do not have the organisational conditions to do so (McCann & Ortega-Argilés, 2015).

How different factors foster technological diversification in countries and regions has been widely scrutinised (Boschma et al., 2015). However, how these factors affect core and peripheral regions differently has lastingly been neglected. Prevailing studies that consider core and peripheral regions have been focusing on, internal and external knowledge networks (Boschma & Iammarino, 2009), the role of entrepreneurial activities (McCann & Ortega-Argilés, 2015), and the quality of government (Rodriguez-Pose & Di Cataldo, 2015). This thesis continues to study the importance of external knowledge networks due to its possibility of active change.

2.3. The Importance of Interregional Linkage

The debate in the diversification of peripheral regions is taken further into the importance of interregional linkages with the emphasis on regions to draw upon external knowledge (Bathelt et al., 2004). The interregional linkages signify that organisations are able to search for knowledge outside of their region which complements the production of knowledge within their region (Balland & Boschma, 2021; Grillitsch & Nilsson, 2015). Therewith, the interregional linkages provide regions with external knowledge and enables regions to diversify into rather distant knowledge from their internal specialisation, and accordingly, preventing regions from a technological lock-in (Asheim & Isaksen, 1997; Isaksen, 2014).

In short, the interregional collaborations as a matter of interregional linkages lead to external knowledge spillovers. Here, external knowledge spillovers is defined as the acquisition of non-local knowledge by organisations as a consequence of collaborative efforts between organisations that locate in distinct regions (Grillitsch & Nilsson, 2015).

Even though previous literature points to the importance of interregional linkage towards diversification in peripheral regions. Two problems arise in the current body of literature.

First, qualitative studies have been focusing on, for example, Onsager et al. (2007) illustrate how external knowledge is complementary to local knowledge for high-technology in small regions in Norway; Fitjar and Rodriguez-Pose (2011) suggest that knowledge production in remote regions in Norway is not a consequence of agglomeration economies but instead of distant social interaction; Fitjar and Rodriguez-Pose (2014) point out that local knowledge is not sufficient for core regions in Norway to diversify; and Rodriguez-Pose and Fitjar (2013) show that promoting external knowledge enhances further generation, diffusion and absorption of knowledge in peripheral regions in Norway. The qualitative studies fail to generalise beyond the sample of individual case studies because qualitative research is not offering a systematic conceptualisation of what the core-periphery dichotomy entails. Therefore, the conceptual definitions and empirical findings in qualitative research are too controversial to conclude on to what extent interregional linkages are important to peripheral regions.

Second, quantitative studies have been focusing on, for example, Grillitsch and Nilsson (2015) observe that external knowledge spillovers complement knowledge production in peripheral regions if the absorptive capacity of a region allows to withdraw external knowledge; De Noni et al. (2017) and De Noni et al. (2018) illustrate that interregional and intraregional linkages in tandem enable both core and peripheral regions to produce knowledge; and Balland and Boschma (2021) show that cognitive proximate interregional linkages enable (particularly peripheral) regions to diversify. The quantitative studies are delimited to a set of indicators to conceptualise interregional linkages on average. Therefore, the prior conceptual definitions of interregional linkage have been too deterministic in terms of to whom regions link and thereby hasn't allowed to differentiate between core and peripheral regions yet.

All in all, this thesis nuances contemporary literature by separating core from peripheral interregional linkages in quantitative means by considering to whom regions link.

According to McCann and Ortega-Argilés (2015), a major aim of EU Cohesion Policy is to let peripheral regions learn from core regions. The question is how to interregional collaboration occurs between core and peripheral regions so that (particularly peripheral) regions benefit from external knowledge spillovers (Iacobucci & Guzzini, 2016). Two issues arise in the current understanding in interregional linkages. First, whether interregional linkages with core regions are more important to those with peripheral regions remains unknown. Second, whether peripheral regions draw upon the external knowledge from core regions, or mainly from peripheral regions, remains unknown. Therefore, this thesis aims to explain the difference on to what extent regions collaborate in terms of their own economic development and the economic development of to whom those regions link.

In this thesis the importance of interregional linkage refers to the extent to which organisations in regions are withdrawing external knowledge through collaboration with organisations in other regions.[1] The importance of interregional linkages is taken as an aggregate of co-inventions between regions (Balland & Boschma, 2021). The novel suggestion in this thesis is to measure interregional linkages with core and peripheral regions separately to study whether external knowledge spillovers affect technological diversification differently over the core-periphery dichotomy. Therewith, the separate conceptualisation of interregional linkages with core and peripheral regions would be able to discern collaboration between regions where knowledge production is low or high.

[1]Note, this thesis neglects the adverse effects of interregional linkage, such as, labour out-migration due to workers that seek fortune, or, exploitation of resources in the periphery by the core (Krugman, 1979).

2.4. Hypotheses

In summary, the principle of relatedness explains that regions tend to specialise into knowledge which is close to their existing knowledge base (Boschma et al., 2015; Dosi, 1982). The organisations prefer to search for nearby knowledge instead of distant knowledge, therefore, the production of knowledge tends to cluster geographically (Balland, 2017; Balland et al., 2020). Therewith, regions face the risk to lock-into nearby knowledge bases, particularly if a region specialises narrowly (Iacobucci & Guzzini, 2016; McCann & Ortega-Argilés, 2015).

The risk to lock-into a narrow specialisation may be overcome by complementing internal knowledge production with external knowledge (McCann & Ortega-Argilés, 2015). Despite the fact that peripheral regions possess weak internal knowledge production, the literature implies any form of external knowledge spillovers complements internal knowledge production (Bathelt et al., 2004). Thus also interregional linkages with peripheral regions do induce knowledge production. The expectation is that both the interregional linkages with core and peripheral regions contribute to diversification in regions. The first hypothesis is:

Hypothesis 1 (H1) *A positive relationship exists between the probability of a region to diversify and the number of interregional linkages with both core and peripheral regions.*

As McCann and Ortega-Argilés (2015) summarise, the six stylised facts explain the core-periphery dichotomy in terms of regions their (in)ability to produce knowledge. Therewith, the fact that interregional linkages enables a region to draw upon external knowledge spillover indicates that interregional linkages with core regions are of greater importance than interregional linkages with peripheral regions. In other words, knowledge production is higher in core regions, and therefore the external knowledge spillovers through interregional linkages with core regions should be beneficial for a region to diversify than through interregional linkages with peripheral regions (McCann & Ortega-Argilés, 2015). In particular, the high knowledge production from linking core regions eases organisations to exploit and recombine the variety in knowledge in comparison to linking peripheral regions (Buenstorf & Klepper, 2009; Klepper, 2007). In terms of path-dependency, interregional linkages with core regions extend the existing knowledge base of a region to a larger variety of knowledge than interregional linkages with peripheral regions (Boschma et al., 2013; Hidalgo et al., 2007), because, the knowledge base in core regions is diverse, and, the knowledge base in peripheral regions is narrow (McCann & Ortega-Argilés, 2015). The second hypothesis is:

Hypothesis 2 (H2) *A stronger relationship exists between the probability of a region to diversify and the number of interregional linkages with core regions compared to the number of interregional linkages with peripheral regions.*

Figure 2.1 depicts the theoretical model. This thesis considers that a set of technological fields is existent within each of the regions. In other words, a region encompasses each technological fields in which the region can specialise or not. The first hypothesis is that the interregional linkages from both core and peripheral regions affect the technological diversification of a technological field positively. The interregional linkages are taken in sum of all technological fields in each region. The second hypothesis is that the interregional linkages from core and peripheral regions are of unequal importance because knowledge production is higher in core regions than in peripheral regions. Nevertheless, how core and peripheral regions link remains open for exploration.

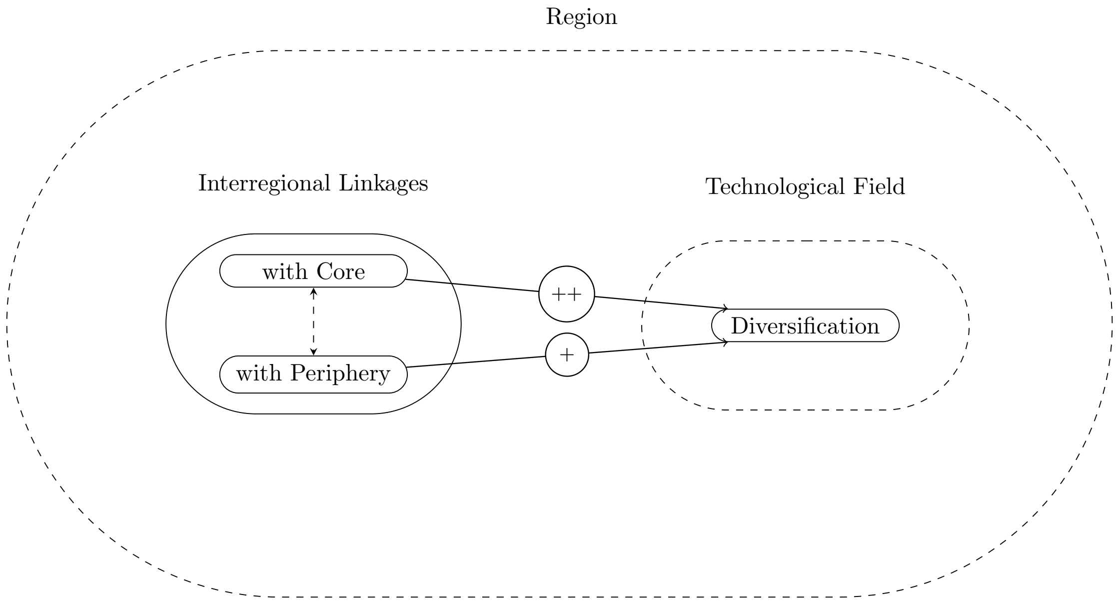


Figure 2.1: Theoretical Model.

The novel stream in literature on the geography of innovation suggests that peripheral regions can be innovative (Eder, 2019; Eder & Trippl, 2019). Therefore, Hypothesis 2 may reverse if a qualitative difference in (all) peripheral regions (in the sample) exist to produce knowledge of certain technological fields. How Hypothesis 2 withstands over different technological fields is an open question that could be answered a posteriori. Potentially, in specific cases, peripheral regions could be the technological frontier, whereas regions are learning from peripheral regions. In other words, instead of considering technological fields equally, here, the suggestion is to allow for variance by considering technological fields also as separate.

3
Methodology

3.1. Research Design

Figure 3.1 depicts the research design as a process. The research starts with gathering the data and turning that into meaningful variables (i.e., data cleaning, imputation, and creation). Subsequently, the research design is split into two approaches. The descriptive analysis of the variables may lead to errors in data cleaning, imputation, and creation. Therefore, the descriptive analysis may feedback in those prior steps to resolve errors.

First, the explanatory approach is to compare the relation between the probability of a region to diversify and the separate effect of interregional linkage with core and peripheral regions. Therefore, a macro-economic perspective is taken to analyse technological diversification in regions over time. In particular, the research employs the entry model by (Boschma et al., 2015). Therefore, the research adheres to an econometric model that operates at the region-technology level. The explanatory approach leads to the confirmation of hypotheses.

Second, the exploratory approach is to propose new insights on technological diversification and interregional collaboration over the core-periphery dichotomy. Therefore, regions are analysed as a network of co-invention. Further, the technology space and metrics on technological fields are explored to investigate how interregional collaboration relates to the production of novel knowledge.

The conclusion and discussion of the thesis follow after the explanatory and exploratory findings. There, the confirmations and the propositions are summarised in terms of the most interesting findings, the recommendations for further research, and the pitfalls of the theory and methodology.

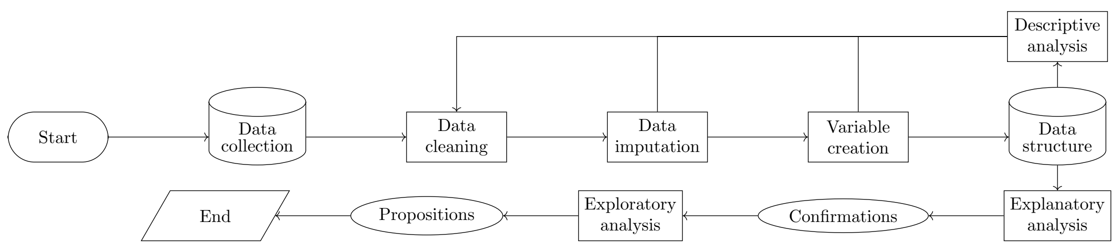


Figure 3.1: Research Process.

3.2. Data & Sample

Within literature of EEG, numerous scientific publications knuckled down onto measuring technological diversification using patent data (Balland et al., 2019; Boschma et al., 2015; Rigby, 2015). Specifically, patents are seen as an extensive source of data to measure knowledge in technological inventions (Whittle & Kogler, 2020). The information within patents as inventions allows to link the counts of technological fields to regions.

First, a technological field among information on inventions is taken by the four-digit Cooperative Patent Classification (CPC4[1]). The CPC4 represents the aggregation of single inventions into an overarching subclass. Second, a region among economic information is taken by the second level of the Nomenclature of Territorial Units for Statistics (NUTS2[2]). The NUTS2 refer to ”basic regions for the application of regional policies” (Eurostat, 2023).

This thesis delimits to Western European regions due to the relation to S3 in the EU Cohesion Policy and the availability of economic and patent data. The consideration of the sample is discussed in Appendix C. The sample consists of ~ 176 NUTS 2 regions over the years $2005 - 2018$ which leads to $\sim 5,000,000$ patents over ~ 650 technological fields by their four-digit CPC. A number of 14 subsequent time frames are taken over the periods as a moving window of five years to aggregate data. This aggregation is necessary to balance the variation in patent application delay, and to increase the data quantity for the robustness of computations (Eck & Waltman, 2009; Steijn, 2021).

To assign patents to regions, the inventor address is taken, and not the applicant address, because the former represents the location of the inventions foremost (Kogler et al., 2017). Therewith, the fractional share of regions and technological fields listed on a patent is taken for more fine-grained data. First, the fractional share by region avoids an overestimation of the regions where headquarters reside. Second, the fractional share by technological field avoids an overestimation of technological fields that frequently list in conjunction of other technological fields (De Rassenfosse et al., 2014).

Two concerns arise in the data. First, patent data is a rich source of data that contains granular data on technological fields e.g., patent texts, citations, inventors, and location. However, an issue in the use of patent data for spatial analysis is the presence of data in core regions, and the absence of data in peripheral regions. Although the research design takes upon a relative approach (Equation 3.1), the absence of data leads to sensitivity in variables (law of small numbers). Thus, the sensitivity leads to under- and overestimation of variables. Accordingly, a threshold is set to a minimum of 50 patents per technological fields at the entire sample to avoid computational issues (Balland et al., 2019) Also a threshold of of ~ 10 patents per period is set to regions, which is explained further in Appendix C. Second, an issue is that missing economic data is mainly from peripheral regions, discarding such observations would be problematic for the validity of the sample. Therefore, a practicality is to impute missing data to involve the periphery, mean imputation is most time-efficient and relatively accurate, or to redefine the sample.

[1] *HTTP: Definition of IPC by World Intellectual Property Organization (WIPO).*
[2] *HTTP: Definition of NUTS by Organization for Economic Cooperation and Development (OECD).*

3.3. Operationalisation

The operationalisation adheres to two independent variables, one dependent variable, and four control variables. An overview of the variable operationalisation is in Appendix A.

3.3.1. The Core-Periphery Dichotomy

The core-periphery dichotomy adheres to the EU Cohesion Policy (McCann & Ortega-Argilés, 2015). The policy assigns regions to objectives. The objectives classify regions as peripheral (lesser developed) regions if the GDP per capita is less than 75% of the EU average, if else, regions classify as core (transition and more developed) regions (Balland & Boschma, 2021).

Furthermore, previous studies take the GDP per capita as an indicator for the core-periphery dichotomy due to the macro-economic relevance to knowledge production (Balland & Boschma, 2021; Rodriguez-Pose & Di Cataldo, 2015; Rodriguez-Pose et al., 2014). On that account, Iammarino et al. (2019) declare that GDP per capita is a good indicator for "education levels, science and technology endowments, infrastructure and institutional quality". Therewith, GDP per capita captures the level of knowledge production over a variety of dimension that aligns directly with the six stylised facts from McCann and Ortega-Argilés (2015).

In summary, the GDP per capita defines the fundamental problem of the core-periphery dichotomy by the EU Cohesion Policy, but also relates scientifically to regional inequality. The dichotomy is taken as binary, and thus not continuous. The binary definition relates directly to the objectives of the EU Cohesion Policy but also eases the interpretation of findings.

Nevertheless, juxtaposing a variety of proxies remains interesting due to lack of consensus in defining the core-periphery dichotomy. In particular, regions are non-linear systems and a lack of one variable often leads to a lack into another variable. However, to allow for variability between regions the comparison of proxies is helpful for robustness checks of a statistical model, and therewith increase the reliability of the results. For example, R&D expenditure per capita, the number of (small) firms, or, the regional innovation scoreboard.

3.3.2. The Technology Space

The principle of relatedness explains how technological fields relate to one each other as existent in the technology space (Boschma et al., 2015). Originally, the technology space derives from the product space by (Hidalgo et al., 2007). Therewith, the technology space is a network representation of technological activities within the economy. The entities (nodes) represent the technological fields and the relationships (edges) represent the relatedness between technological fields. In the technology space, two technological fields are observed as related if they co-occur on a patent. In this thesis the technology space as a visual representation of the network is useful to understand how technological fields arise differently within core and peripheral regions, but also to grasp how interregional collaboration may relate to the relatedness between technological fields. Therewith, the assumption is that the technological fields at the periphery of the technology space are least related to the technological activities within the sample (Fleming & Sorenson, 2001). Thereby, the unrelated technological fields are novel in the economy (Balland, 2017).

3.3.3. Interregional Linkages

The interregional linkages are taken by the number of co-inventors of patents that reside in various regions as direct extension to Balland and Boschma (2021). The residence of inventors represent that the regions of co-inventors (within organisations) embody the knowledge from patented inventions. Therewith, co-invention signifies that organisations collaborate between regions and cause external knowledge spillovers. All in all, the measure of interregional linkages is defined as the total centrality degree of co-inventors in a region as an undirected network (Le Gallo & Plunket, 2020; Whittle et al., 2020).

This thesis delineates interregional linkages with core regions $ILC_{r,t}$ from interregional linkages with peripheral regions $ILP_{r,t}$. Thus, the variables $ILC_{r,t}$ and $ILP_{r,t}$ are the quantities of inventors in a region r having an inventors in remaining regions s that classify as either a core or peripheral region. The measures of interregional linkages are taken for each region r at time t specifically as an aggregation of all technological fields.[3] Thus, no distinction is made in external knowledge spillovers between technological fields i.

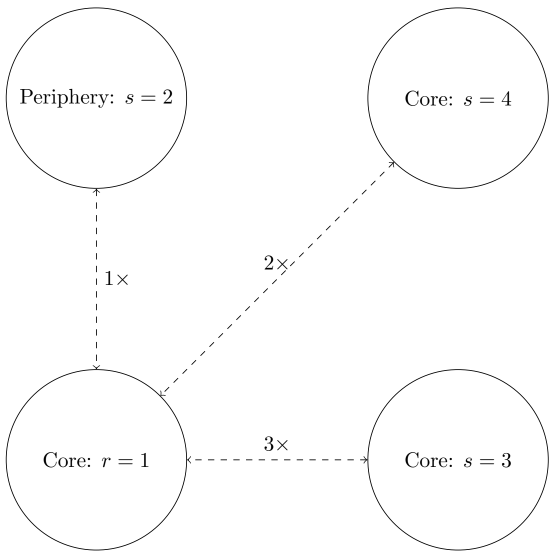


Figure 3.2: The operationalisation of interregional linkages with core and peripheral regions.

For example, in Figure 3.2 region $r = 1$ links to regions $s = 2$ once, $s = 3$ thrice, and $s = 4$ twice within time frame t. Therewith, region $s = 2$ is a peripheral region and $s = 3$ and $s = 4$ are core regions. Then, for $r = 1$ the number of interregional linkages with core regions is $ILC_{1,t} = 2 + 3 = 5$ and those with peripheral regions is $ILP_{1,t} = 1$. On top of the number of interregional linkages with core and peripheral regions also the total number of interregional linkages $ILT_{r,t} = ILC_{r,t} + ILP_{r,t}$ is of interest as a baseline to compare the statistical effects of the interregional linkages with core and peripheral regions.

[3]Note, no data about the core-periphery dichotomy is available (at hand) outside of the sample. Thus, interregional linkages do not exist with regions in other continents for example.

3.3.4. Entry of a Technological Field

The dependent variable technological diversification is operationalised by the entry of a technological field in a region. The entry model measures technological diversification in regions and allows for path-dependency statistically at the region-technology level. The entry model considers a region to specialise into a technological field if the region relatively has a higher number of patents in a technological field in comparison to the reference sample. Therefore, the entry model is an appropriate representation of technological diversification because the model captures in a simplified manner whether a region does or does not specialise into a technological field that is new to the region (Boschma et al., 2015).

The entry of a technological field is defined as the Relative Technological Advantage $RTA_{i,r,t}$ which signifies whether a region r specialises in a technological field i at time t (Balland, 2016) in comparison to the reference sample:

$$RTA_{i,r,t} \begin{cases} 1, & \text{if } \frac{p_{t,r,i}}{\sum_i p_{t,r,i}} > \frac{\sum_r p_{t,r,i}}{\sum_r \sum_i p_{t,r,i}} \\ 0, & \text{otherwise} \end{cases} \tag{3.1}$$

thus $RTA_{i,r,t} = 1$ if a region r yields relatively more patents p in a technological field i at time t. In other words, the indicator is a binary variable which is true if the regional share in a technological field is greater than the reference sample (the EU). The time frame in which a region specialises in a technological field i.e., the time frame in which $RTA_{i,r,t} = 0$ turns into $RTA_{i,r,t} = 1$, is the entry of a technological field $entry_{i,r,t}$. The variable takes no value (*NA*) if $RTA_{i,r,t} = RTA_{i,r,t-1} = 1$, which means a region specialises in a technological field already.

All in all, the entry model captures whether a region specialises into a new technological field regardless of the quantity of patents in a region or technological field. Further, no inherent bias exist towards the propensity to patent in high-technology industries and core regions. Therefore, the entry model is especially relevant to study technological diversification in peripheral due to the ability to overcome the general bias in patent data.

Within the entry model it is crucial to control for the relatedness density i.e., the knowledge gap within a region towards diversifying into a technological field, to take into account the existing knowledge base of a region (Boschma et al., 2015). Notable is that the entry model measures horizontal development as of the width of technological diversification, but neglects vertical development as of the depth of technological specialisation. Despite that technological diversification is important for peripheral regions to broaden their economic activities, peripheral regions may rely more on technological specialisations due to their path-dependency (Hassink & Gong, 2019). How the relation between technological diversification and specialisation should be handled into statistical modelling is unresearched and therefore is not taken into account in this thesis.

3.3.5. Relatedness Density

An essential control variable is the relatedness density, which is an indicator of the existing knowledge in a region. The relatedness density signifies the nearness of a technological field to the portfolio of a region (Boschma et al., 2015). A prior step is to compute the relatedness $\varphi_{i,j,t}$ between pairs of technological fields i and j on patent applications (Balland, 2016):

$$\varphi_{i,j,t} = \min\{P(\mathit{RTA}_{i,t} > 1 | \mathit{RTA}_{j,t} > 1), (\mathit{RTA}_{j,t} > 1 | \mathit{RTA}_{i,t} > 1)\} \tag{3.2}$$

herewith, the relatedness $\varphi_{i,j,t}$ is the minimum of the pair-wise conditional probabilities of regions that specialise into a technological field i as well as a technological field j at time t. The approach is to count the co-occurrence between the primary listed technological fields i and j within a region r, and subsequently standardising the count by the total number of patents in technological fields i and j. The co-occurrences are normalised through cosine similarity indexing to preserve the independence assumption between entities (Eck & Waltman, 2009). The subsequent step is to compute the relatedness density $\mathit{RD}_{i,r,t}$ of technological field i in region r for a region to diversify (Balland, 2016):

$$\mathit{RD}_{i,r,t} = \frac{\sum_{j \in r, j \neq i} \varphi_{i,j}}{\sum_{j \neq i} \varphi_{i,j}} \cdot 100\% \tag{3.3}$$

so, the density in the production of knowledge in technological field i for region r at time t is build upon $\varphi_{i,j}$ as of the relatedness of technological field i to the remaining technological fields j in which a region has an $\mathit{RTA}_{i,r,t} = 1$ relative to the reference sample (the EU). The relatedness density $\mathit{RD}_{i,r,t}$ yields a minimum value of 0% if no technological fields relate to technological field i in region r, and yields a maximum value of 100% if all remaining technological fields related to technological field i co-occur in region r at time t.

3.3.6. Covariates

Next to the independent and dependent variables four control variables are taken in line with Balland and Boschma (2021). First, the expectation is that the level of population size $PopSize_{r,t}$ (log) yields a positive effect on the probability of a region to diversify because this depicts the effect of total variety of knowledge in a region. Second, the expectation is that the level of GDP per capita $GDP_{r,t}$ yields a positive effect on the probability of a region to diversify because this depicts the effect of economic development (i.e., education levels, science and technology endowments, infrastructure and institutional quality). Third, the inclusion of periodical fixed-effects α_t adjust for variation in technological change over time. Fourth, the interaction between absorptive capacity $AC_{r,t}$ and interregional linkages is of interest because regions are limited to the extent they are able to absorb external knowledge spillovers. The absorptive capacity is taken by the gross R&D expenditures of a region (Lau & Lo, 2015). Fifth, the population density is taken into account to control for the effect of agglomeration economies whereas dense population lead to the exchange of knowledge. Further, control variables are taken by fixed-effects of countries ϕ_c and technological classes ψ_i. The entire description on control variables follows in the empirical findings.

3.4. Analytical Strategy

The research design revolves around both an explanatory and an exploratory approach. But, the research kicks-off with a descriptive analysis to assure the quality of variables.

3.4.1. Explanatory Analysis

This thesis revolves around the two independent variables of interregional linkages with core and peripheral regions and one dependent variable of the entry of a technological field in a region. This thesis investigates how these two independent variables (H1) affect the dependent variable, and how these independent variables are different from each other in their effects on the dependent variable (H2). Therefore, a single linear probability model is of interest to compare the coefficients of the two independent variables. The equation to estimate the linear probability model is:

$$\begin{aligned} Entry_{i,r,t} = \beta_1 \cdot ILC_{i,r,t-1} + \beta_2 \cdot ILP_{i,r,t-1} + \beta_3 \cdot RD_{i,r,t-1} \\ + \beta_4 \cdot GDP_{r,t-1} + \beta_5 \cdot PopSize_{r,t-1} + \alpha_t + \epsilon_{i,r,t} \end{aligned} \tag{3.4}$$

The first hypothesis (H1) is simply analysing the significance of the coefficients. The second hypothesis is analysis the regression model as stepwise variable selection. The straightforward way is to first include the variable $ILP_{i,r,t}$, and then the variable $ILC_{i,r,t}$ separately and concurrently. Both coefficients should be significant, however, if the former coefficient loses significance or turns in sign after adding the latter coefficient that would imply the interregional linkages with core regions absorbs the variance of those with peripheral regions. In other words, linking to the core may be more determinant than linking with the periphery. The same analytical process is done to compare the effect of $ILP_{i,r,t}$ and $ILC_{i,r,t}$ to the base line of the total number of interregional linkages $ILT_{i,r,t}$.

To investigate the core-periphery dichotomy further, the binary distinction at 75% of the average GDP per capita is taken as a parameter to create further binary distinctions between $50\% - 150\%$ of the average GDP per capita. Therewith, the effect of the interregional linkages with core $ILC_{i,r,t}$ and with peripheral $ILP_{i,r,t}$ regions are scrutinised over this parameter.

Further, stratified sampling are done by running the linear probability model for core and peripheral regions separately. By this means, whether interregional linkages with core and peripheral regions is more important for peripheral regions can be analysed. Also, robustness checks are done by adding more variables i.e., absorptive capacity $AC_{r,t}$, population density $PopDens_{r,t}$, and fixed-effects of countries ϕ_c and technological classes ψ_i. These fixed-effects are set up by disaggregating the labels of regions and technological fields into 2-digit labels. Further, robustness checks are done by stratifying the sample over two equal samples in time and applying a logistic regression model instead of a linear regression model.

3.4.2. Exploratory Analysis

To take the analysis of technological diversification the exploratory analysis investigates the regions and technological fields further. At the regional level, the collaboration between core and peripheral regions is scrutinised through social network analysis (SNA). At the technological level, several measures of the technology space are investigated.

First, the regional analysis puts regions in quadrants by the dimensions of economic development and the relative number of entries in regions. These quadrants are useful to depict over- and underachieving regions at a geographical map and as a social network of regions. Here, the social network consists of regions as entities (nodes) and co-inventions as relationships (edges). The SNA is interesting to explore as a visualisation but also according to centrality measures. The degree, eigenvector and betweeness are of interest. The degree centrality is the total number of interregional linkages and measures the direct extent to which external knowledge spillovers are available to a region. The eigenvector centrality is the extent to which a region's neighbours connect and measures the indirect extent to which external knowledge spillovers are available to a region. The betweenness centrality is the shortest path to other regions and measures the extent to which a region is a gatekeeper, or a bridge, of knowledge flow between regions. The centrality measures are analysed per quadrant by taking the average.

To understand whether interregional linkages tend to mix between core and peripheral regions, or not, the measure of assortative mixing by enumerative characteristics is applied (Newman, 2002). The assortativity coefficients indicates homophily versus heterophily in a social network of co-invention. The social network of interregional linkage consists of regions as entities (nodes) and co-inventions as relationships (edges). The explanation of assortative mixing is in Appendix B. The assortativity coefficient is computed for each time frame of five years over the period $2005 - 2018$. The final equation is:

$$-1 \leqslant r = \frac{Q}{Q_{max}} = \frac{A_E - E_E}{m - E_E} \leqslant 1. \tag{3.5}$$

here the assortativity coefficient r is an outcome of the actual linkages between regions of the same objective A_E and the expected linkages between regions of the same objective E_E in comparison to the total number of linkages in the sample m. Here, $r = -100\%$ if all co-invention is between core and peripheral regions, but not between core regions themselves and peripheral regions themselves, and $r = 100\%$ if all co-invention is between core regions themselves and peripheral regions themselves, but not between core and peripheral regions. Lastly, $r = 0\%$ if co-invention distributes as if the network is random.

Second, technological analysis puts technological fields in quadrants by the dimensions of the assortativity coefficient for each technological field and the difference in entries over the core-periphery dichotomy. Here, the assortativity coefficient for each technological field refers to the extent to which core and peripheral regions collaborate (or not) over that specific technological field. The difference in entries over the core-periphery dichotomy indicates whether a technological field relatively arises more in core or in peripheral regions. These quadrants are useful to analysis the potential in collaboration between core and peripheral regions in conjunction with the competitive advantage that core and peripheral regions take.

The technology space is explored further with the measures of technological ubiquity and novelty. The technological ubiquity measures the number of regions in which a technological field is present in terms of Relative Technological Advantage $RTA_{i,r,t}$. The technological ubiquity describes whether a technological field is unique (or not) and is important to understand whether core or peripheral regions take a competitive advantage over such ubiquity (Balland & Rigby, 2017). Further, the technological novelty here is taken by the mean relatedness of a technological field to other technological fields (Balland, 2016). In terms of the recombination of inventions, a low mean relatedness refers to that a technological field is novel due to few synthesis of existing technological fields, but is also risky due to the uncertainty of a successful outcome i.e., whether a recombination of inventions reaches the implementation to society (Fleming & Sorenson, 2001). Therewith, a high mean relatedness refers to low novelty and low risk (Balland, 2016). The technological ubiquity and novelty are investigated over the quadrants that put the technological fields into perspective of interregional collaboration between core and peripheral regions.

3.5. Quality Criteria

The quality of this thesis revolves around reliability and validity (Bryman, 2016). First, the external validity delimits to the sample of the EU member states, which is relatively the high-end in the spectrum of technological change and the periphery, therefore, technology is considerably more present in comparison to the worldwide spectrum. How the technological diversification persists in developing (i.e., the low-end in the spectrum of technological development) regions is abdicated. Nevertheless, this thesis is relevant for S3 policy and economic cohesion of the EU. Second, the internal validity delimits to the patentability of inventions in the sample of the EU member states. As Pugh and Dubois (2021) state "it's all relative" in peripheral regions. Therewith, through missing data imputation and the approach of *RTA* these issues resolve, however, *RTA* is a sensitive to a low number of patents in a region, such observations must be removed. Third, external reliability delimits to the operationalisation of proxies, nonetheless, performing robustness checks over multiple proxies improves external reliability. Fourth, internal reliability delimits to the comparability between time periods and technological fields, improvements are done through controlling for time fixed-effects and adjusting for technological fields as separate.

Furthermore, problems arise using regions as units. First, NUTS2 are statistical units for spatial analysis, however, regions remain difficult to compare i.e., regions are not a systematic unit of comparison. Second, 'correlation is no causation' which is even more true for regions as non-linear systems. Therefore, controlling for variables for causal inference is dubious, and the results are rather correlational. Therefore, complete certainty in the interpretation of coefficients does not exist, and thus, relationships are discussed in terms of associations but not in causal terms.

4
Empirical Findings

4.1. Descriptive Analysis

As first part of the empirical findings, the descriptive analysis creates an understanding in the data and puts the variables into their context before proceeding with the explanatory and exploratory analysis. The descriptive analysis starts with Figure 4.1 to depict core and peripheral regions by economic development over time.[1] Unless specified otherwise, the core-periphery dichotomy is split at 75% of the sample mean GDP per capita at a given year.

Figure 4.1 indicates that the core-periphery dichotomy shifts to the Southern countries in Western Europe over time. In Figure 4.1a in 2000 (North) Germany (DE) has more peripheral regions than in Figure 4.1d in 2015. In contrast, in Figure 4.1a in 2000 France (FR) has fewer peripheral regions than in Figure 4.1d in 2015. Further, in Figure 4.1b in 2005 and in Figure 4.1c in 2010 (North-East) Spain (ES) has fewer peripheral regions than in in Figure 4.1a in 2000 and in Figure 4.1d in 2015. Lastly, the number of peripheral regions has risen in (North) Italy (IT) from 2010 and on-wards as in Figure 4.1c. Apart from a small shift in peripheral regions, a set of core regions comprise core countries that do not turn peripheral at any time in the sample. Swiss (CH), Denmark (DK), Finland (FI), Luxembourg (LU), Netherlands (NL), Norway (NO), Sweden (SE) do not comprise of any peripheral region at any time within this sample.

In general, peripheral regions are geographically proximate to peripheral regions (and vice versa for core regions).[2] Further, since peripheral regions reside in peripheral countries, peripheral regions would be culturally proximate to peripheral regions as well (Boschma, 2005). One interesting notice is that none of the capital regions except Área Metropolitana de Lisboa is peripheral within this sample which emphasises economic growth within capitals.

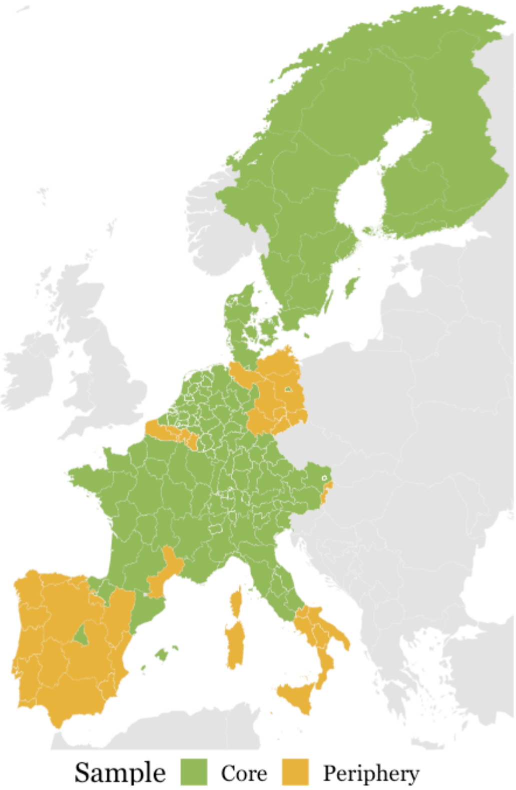
The definition of the core-periphery dichotomy is somewhat sensitive to those regions that reside at the boundary of being a core or a peripheral region. However, as Table 4.1 indicates, the core-periphery dichotomy remains roughly stable over time considering the number of core and peripheral regions in Figure 4.1.

Table 4.1: The number of core and peripheral regions over time.

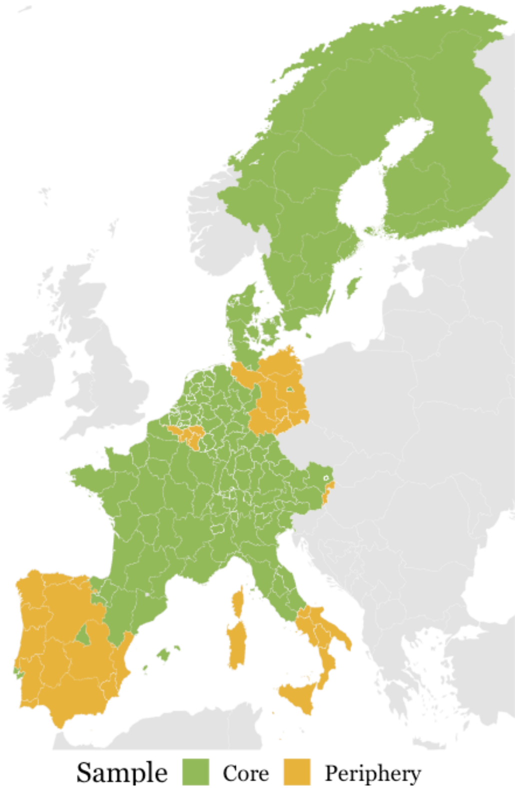
	2000	2005	2010	2015
Number of core regions	125	129	128	124
Number of peripheral regions	46	42	43	47

[1]Note, Appendix C discusses the reselection of the sample to create a coherent sample of Western European regions and countries (for which economic and patent data is available).

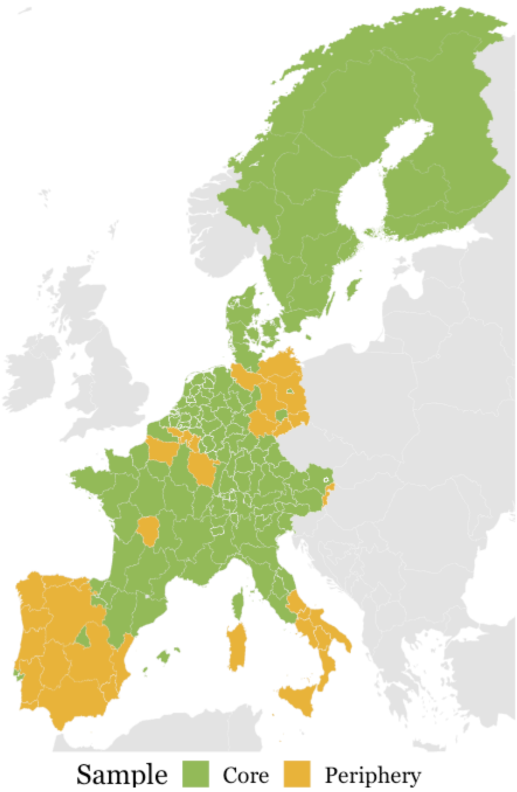
[2]The geographical proximity between core and peripheral regions underling indicates spatial autocorrelation may be necessary to accommodate for spatial dependence in regression models.



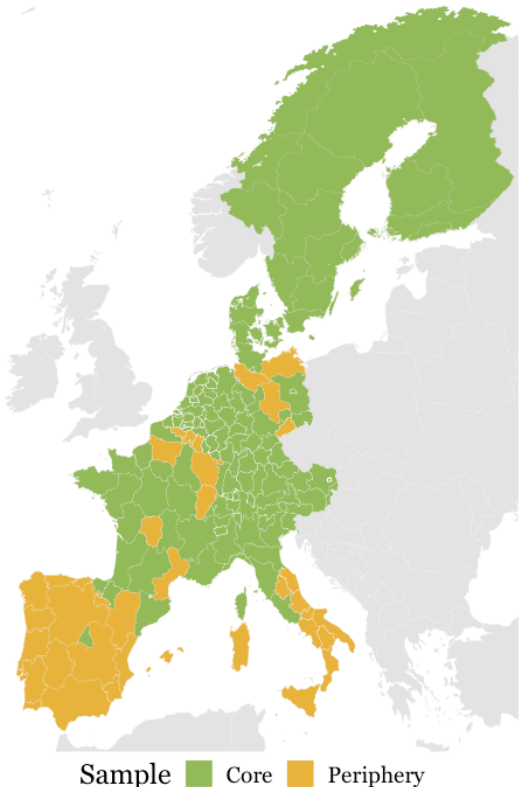
(a) The year 2000.



(b) The year 2005.



(c) The year 2010.



(d) The year 2015.

Figure 4.1: The core-periphery dichotomy relative to Western European regions over time.

The following analysis concerns the descriptive statistics of all the variables in this thesis. Table 4.2 summarises the variables for core and peripheral regions separately. The descriptive statistics are useful to interpret the difference between core and peripheral regions their mean and their maximum values of variables.

Table 4.2: Descriptive statistics stratified over the core-periphery dichotomy.

Statistic	N	Mean	St. Dev.	Min	Max
Core					
Entry	828,518	0.136	0.343	0	1
Relatedness Density	1,101,325	27.487	17.702	0	100
Linkages Core	1,101,325	1,768	2,162	5	11,278
Linkages Periphery	1,101,325	107	189	0	1,953
GDP per Capita	1,101,325	36,877	12,423	21,500	98,800
Population Size	1,101,325	1,972,447	1,799,714	26,530	12,213,447
Population Density	1,101,325	384	831	3	7,472
R&D Expenditure	1,040,586	3,787	5,633	1	18,664
Periphery					
Entry	328,858	0.073	0.261	0	1
Relatedness Density	383,682	18.379	16.062	0	100
Linkages Core	383,682	307	535	0	3,352
Linkages Periphery	383,682	47	67	0	425
GDP per Capita	383,682	19,976	3,847	5,300	27,600
Population Size	383,682	1,777,506	1,673,166	84,708	8,410,095
Population Density	383,682	182	369	2	6,059
R&D Expenditure	383,682	2,292	5,195	0	18,664

First, the mean of the entry of a new specialisation in a technological field is higher in core (0.136) than in peripheral (0.073) regions. Thus, in this sample of Western European regions, the economic dichotomy disseminates into a technological dichotomy as well. This indicates that the core regions are leading, whereas the peripheral regions are lagging, in technological diversification (Hassink & Gong, 2019; McCann & Ortega-Argilés, 2015).

Second, the mean of the relatedness density is higher in core (27.487%) than in peripheral (18.379%) regions. The difference emphasises the path-dependency of technological diversification over the core-periphery dichotomy. Since core regions specialise into a broader set of new technological fields, the relatedness density grows higher in core regions in comparison to peripheral regions, and as a consequence, the entry of new specialisations in technological fields reflects back upon the relatedness density as a virtuous cycle for core regions (Whittle & Kogler, 2020). In other words, the knowledge base in core regions grows wider due to the relatedness to new technological fields whilst peripheral regions tend to lock-in into a narrow knowledge base due to the lack of relatedness to new technological fields.

Third, and not surprisingly, also the mean of the interregional linkages of both the types (i.e., with core and with peripheral regions) are higher in core ($1,768$ & 107) than in peripheral (307 & 47) regions. The statistics indicate that core regions are better connected within the sample than peripheral regions are. Therewith, the difference between core and peripheral regions is less explicit in the mean of the interregional linkages with peripheral regions ($107/47 \approx 2$) in comparison to the mean of the interregional linkages with core regions ($1,768/307 \approx 6$). The results indicate that peripheral regions are rather proximate to other peripheral regions in contrast to core regions in terms of co-invention. From another perspective, the results indicate that the core regions dominantly collaborate mainly to other core regions ($1,768/107 \approx 17$) in comparison to peripheral regions ($307/47 \approx 6$). The region with the most interregional linkages with core regions ($11,278$) is Karlsruhe in 2009. The the region with the most interregional linkages with peripheral regions ($1,953$) is Berlin in 2013. These two regions are both core regions in Germany.

Fourth, the mean population size is slightly lower in peripheral ($1,777,506$) than in core ($1,972,447$) regions. Further, the maximum population size is lower in peripheral ($8,410,095$) than in core ($12,213,447$) regions. Thus, in general, no notable differences exist in the population size between core and peripheral regions. Nevertheless, the largest population resides in a core region, which is Île de France in 2018.

Fifth, the mean population density (inhabitants per km^2) is lower in peripheral (384) than in core (182) regions. Likely, peripheral regions comprise less cities and more rural areas in comparison to core regions. The difference between core and peripheral regions in the minimum and maximum population density is less apparent. The region with the highest population density ($7,472$) is Brussels, which is a core region.

Sixth, the mean R&D expenditure (in million Euro) is lower in peripheral ($2,292$) than in core regions ($3,787$). In line with the core-periphery dichotomy in economic development, the core regions spend more than peripheral regions on R&D. The difference between core and peripheral regions in the minimum and maximum R&D expenditure is negligible.

As a next step, Table 4.3 comprises the correlations between all the variables in this thesis to investigate generic associations. The interpretation of the Pearson's correlation coefficients (i.e., weak, moderate, and strong associations) for political science by Akoglu (2018) is taken to highlight several associations between variables.

First, a set of correlations yields moderate associations ($0.4 \leqslant r < 0.7$). In line with prior expectations, a moderate association exists between the interregional linkages with core and the interregional linkages with peripheral regions within the sample. Thus, according to the extent to which a region is outward looking, a moderate association exists to linking with both core and peripheral regions. Besides, according to the population size, a moderate association exists to linking more with both core and peripheral regions. Straightforwardly, a large population size leads to more inventors and thus more interregional linkages in general. Those moderate associations may cause multicollinearity which are problematic in regression models. The separate consideration of these variables in regression model is of importance to prevent undesirable results in the interpretation of variable coefficients.

Second, a set of correlations yields weak associations ($0.3 \leqslant r < 0.4$). The GDP per capita yields a weak association to interregional linkages with core regions, but not with peripheral regions. In other words, richer regions tend to have more links to other rich regions, but not as much to poorer regions. Next to this, the population density of a region has a weak association to interregional linkages with peripheral regions, but not with core regions. In other words, high density regions associate with linking to poor regions, but not so much to richer regions. Contradictory, the population density of a region has a weak association with GDP per capita. Likely, densely populated regions comprise larger cities which tend to be the relatively richer regions in the sample. Nevertheless, the associations are weak and thus are deceptive because these weak associations do not deem to be significant relationships.

Table 4.3 does not indicate strong associations ($r \geqslant 0.7$) between variables in this thesis. Nonetheless, further analysis of associations through regression models using multivariate statistics would lead to substantiated findings that suffer less from the effect of spurious correlations. Specifically, correlations do not indicate true relationships, therefore, regression models (and theory) are necessary to exclude confounding causes. All in all, one should not conclude on relationships on basis of correlations.

Table 4.3: Correlation between variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Entry	1							
(2) Relatedness Density	0.204	1						
(3) Linkages Core	0.058	0.174	1					
(4) Linkages Periphery	0.043	0.121	0.488	1				
(5) GDP per Capita	0.055	0.124	0.336	0.201	1			
(6) Population Size	0.063	0.199	0.392	0.405	0.025	1		
(7) Population Density	0.016	0.029	0.168	0.379	0.347	0.103	1	
(8) R&D Expenditure	0.000	-0.054	0.077	0.075	-0.119	0.193	-0.037	1

Figure 4.2 indicates the percentage mean entries (versus non-entries) of new specialisations in any technological field relative to all technological fields over regions between the core-periphery dichotomy. The figure indicates, alike Table 4.2, that the entry of a technological field is less likely in peripheral regions than in core regions. However, when considering the trend (moving-window of five years) over time[3], core regions remain rather stable in technological diversification (+1%), whilst peripheral regions show an increasing amount of technological diversification (+3%). The increasing trend in peripheral regions seems to flatten at the end of the time frame. Nevertheless, considering the standard deviations around the means, not every peripheral region has less entries than some of the core regions do.

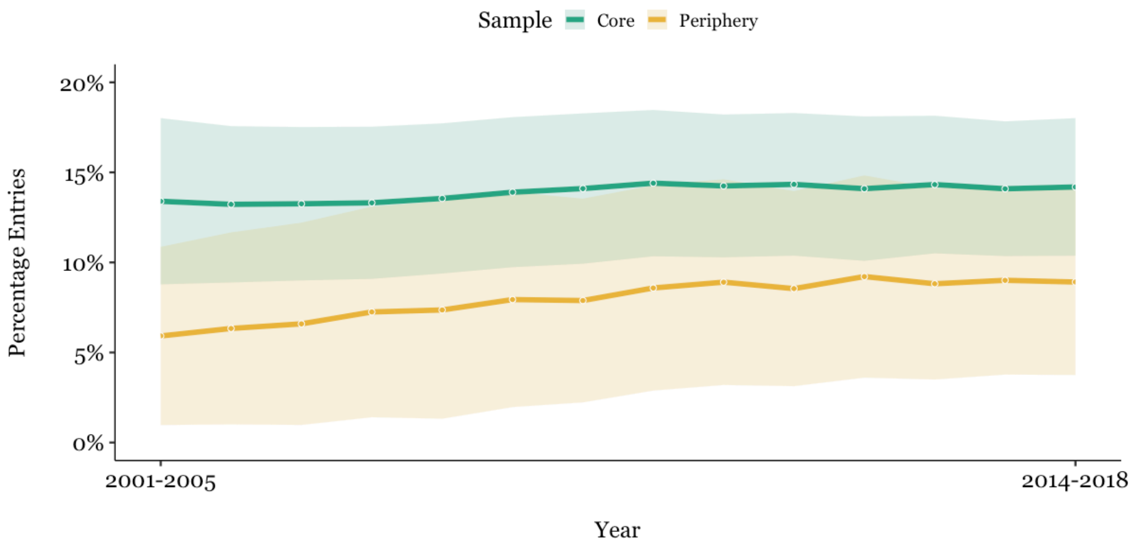


Figure 4.2: The mean percentage of total entries in technological fields in regions over time between the core-periphery dichotomy. The shaded ribbon indicates the standard deviation around the mean. The data is grouped corresponding to whether the region is core or peripheral.

Putting the results of Table 4.2 and Figure 4.2 in line with theory, the magnitude of knowledge production between core and peripheral regions defined by economic development separates the extent of technological diversification (Balland & Boschma, 2021). In other words, the core regions are leading regions whereas the peripheral regions are lagging regions because the core regions run ahead from peripheral regions in technological diversification (McCann & Ortega-Argilés, 2015). As Hassink and Gong (2019) state, peripheral regions are followers of knowledge production of core regions. Accordingly, the core regions define the technological frontier, and thus the peripheral regions require to learn from novel knowledge through external knowledge spillovers from, especially, the core regions. Therewith, the trend indicates that peripheral regions attempt to catch-up, but, whether this trend continues, and thus, whether the difference in technological diversification between core and peripheral regions may even out over the long run is unpredictable. Interestingly, whilst the gap in economic development has grown, the gap in technological diversification has shrunk. Yet, peripheral regions may be leading over core regions in some aspects whilst overall they do not.

[3]Note, the year 2005 indicates the aggregation of patents over $2001-2005$ and the year 2018 that of $2014-2018$. In total, the moving-window consists of 14 periods of five years over $2005-2018$.

4.2. Explanatory Analysis

The regression analysis consists of 621 technological fields within 172 regions in order to estimate the separate effect of interregional linkages with core and with peripheral regions on technological diversification. The analysis consists of a central model and multiple robustness checks to confirm whether the results withstand under different circumstances. All regression tables are relevant to test the two hypotheses.

Table 4.4 is the central model and involves the variables relatedness density, GDP per capita, population size, and time-fixed effects as a baseline of control variables. In short, all variables have a significant positive effect on the entry of a technological field in a region. Note, the variables are not standardised, therefore variables of different units are not comparable among each other. Nevertheless, the goal is to investigate the variables of interregional linkages with core and with peripheral regions. These two variables are of an identical measure, and thus are directly comparable. The variable R&D expenditure, interaction effects, and country & technology fixed-effects are excluded from the central model to have a directly comparable model (apart from the sample) to the model by Balland and Boschma (2021). Notable is that the magnitudes of all variable regression coefficients are identical to Balland and Boschma (2021). Nevertheless, in this thesis the interregional linkages with core and peripheral regions are taken separately to investigate the core-periphery dichotomy.

First, relatedness density has a positive effect on the probability of a region to diversify into a new technological field.[4] Therefore, regions diversify into technological fields that are nearby to their existing knowledge base. This result confirms the importance of path-dependency for a region to diversify (Balland et al., 2019; Boschma et al., 2015; Rigby, 2015).

Second, GDP per capita has a positive, yet negligible, effect on the probability of a region to diversify into a new technological field[4]. Thus, the level of economic development contributes to a region to diversify (Petralia et al., 2017), but is almost nil, especially when considering the coefficient in conjunction with the range of the variable as in Table 4.2, which is $\sim 10,000 - 100,000$ Euro per inhabitant, and is relatively low in comparison to other variables.

Third, population size has a positive effect on the probability of a region to diversify into a new technological field[4]. The population size leads to more inventors and thus more knowledge production (Boschma et al., 2015). The population size has a large effect when considering the coefficient in conjunction with the range of the variable as in Table 4.2, which is $\sim 1,000,000 - 10,000,000$ people and is relatively high in comparison to other variables.

Fourth, the total number of interregional linkages has a positive effect on the probability of a region to diversify into a new technological field in model 1. Therefore, the external knowledge spillovers indeed contribute to a region to diversify (Balland & Boschma, 2021). However, when involving the number of interregional linkages with core and peripheral regions, the total number of interregional linkages loses its positive effect. The loss of the positive effect of this variable is discussed further in the following subsections in line with the hypotheses.

[4]The results remain stable considering all seven models in Table 4.4.

4.2.1. Effects of Interregional Linkages with Core and Peripheral Regions

In Table 4.4 model 1, 2, and 3 indicate that any variable of interregional linkages, may it be the total, or, solely with core and peripheral regions, has a positive effect on the probability of a region to diversify into a new technological field when considered individually. Therefore, the risk to lock-into a narrow specialisation may be overcome by complementing internal knowledge production with external knowledge production from both core and peripheral regions (McCann & Ortega-Argilés, 2015). The results imply that both external knowledge spillovers from core and peripheral regions complement internal knowledge production (Bathelt et al., 2004). All in all, both interregional linkages with core and peripheral regions seem to contribute to knowledge production. The results in model 2 and 3 together with the perception of literature provides evidence to accept H1. Thus, a positive relationship exists between the probability of a region to diversify and the number of interregional linkages with both core and peripheral regions. However, as shown in Table 4.3, a note should be made that the two variables of interregional linkages with core and peripheral regions correlate. Thus, the positive effect of interregional linkages with peripheral regions could be attributable to the effect of interregional linkages with core regions. The following hypotheses explores this doubt further.

In model 6 and 7 of Table 4.4, when considering multiple variables of interregional linkages into a single model, the effect of the total number of interregional linkages and those with peripheral regions diminish. In particular, the effects of those variables turn negative whilst the effect of interregional linkages with core regions remains positive. The fact that the effect of two out of three variables of an identical measure change drastically is due to multicollinearity between independent variables. The variance inflation factor (VIF) explains statistically whether multicollinearity occurs. In model 4-7, between all the variables of interregional linkages considerable collinearity is present due to a $VIF > 2.5$ (Johnston et al., 2018). In this case, the effect of interregional linkages with core regions takes over the explainable variance of the other variables, and thus explains the entry of technological fields best. This collinearity is discussed further along the model evaluation & diagnostics.

Furthermore, McCann and Ortega-Argilés (2015) substantiate theoretically why the loss of effects occur. Namely, the six stylised facts explain the core-periphery in terms of regions their (in)ability to produce knowledge. Here, the results indicate that interregional linkages with core regions are of greater importance because knowledge production is higher in core regions. Therewith, an interregional linkages with a core region weights higher than interregional linkages in general and interregional linkages with peripheral regions because higher knowledge production leads to more knowledge spillovers between regions. In other words, the cost of an interregional linkages with a peripheral region is higher than the cost of an interregional linkages with a core region due to an inherent different quality of external knowledge spillovers. The results in model 6 together with the perception of literature provides evidence to accept H2. Thus, a stronger relationship exists between the probability of a region to diversify and the quantity of interregional linkages with core regions compared to the number of interregional linkages with peripheral regions.

Table 4.4: The central technological diversification model i.e., entry, or not, of a new specialisation in a technological field in a region. All the models refer to Equation 3.4 (OLS) and include time fixed-effects α_t Balland and Boschma (2021). The models 2 and 3 are relevant to test H1, and the models 6 is relevant to test H2.

	Dependent variable:						
	$Entry_{i,r,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relatedness Density	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
GDP per Capita	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Population Size	0.002***	0.002***	0.010***	0.003***	0.004***	0.003***	0.003***
(log)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0005)	(0.0005)	(0.0005)
Linkages Total	0.009***			−0.015***	0.012***		−0.009***
(log)	(0.0002)			(0.002)	(0.0003)		(0.002)
Linkages Core		0.009***		0.024***		0.011***	0.019***
(log)		(0.0002)		(0.002)		(0.0003)	(0.002)
Linkages Periphery			0.003***		−0.003***	−0.002***	−0.002***
(log)			(0.0002)		(0.0003)	(0.0003)	(0.0003)
Period	Included	Included	Included	Included	Included	Included	Included
Constant	−0.060***	−0.057***	−0.142***	−0.064***	−0.082***	−0.075***	−0.072***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	1,157,376	1,157,376	1,157,376	1,157,376	1,157,376	1,157,376	1,157,376
R^2	0.045	0.045	0.044	0.045	0.045	0.045	0.045
Residual Std. Error	0.316	0.315	0.316	0.315	0.315	0.315	0.315
F Statistic	3,225.501***	3,234.520***	3,143.085***	3,059.680***	3,054.954***	2,899.709***	2,899.709***
Note:							*p<0.1; **p<0.05; ***p<0.01

4.2.2. Parameterisation of the Core-Periphery Dichotomy

Thus far the core-periphery dichotomy is split at 75% of the average GDP per capita. To investigate the hypotheses further, the core-periphery dichotomy is split at multiple positions. The core-periphery dichotomy and variable creation is re-split in a range of 50%-150% in steps of 5% to investigate how the dichotomy in external knowledge spillovers exists. For each recomputation of the core-periphery dichotomy the regression is repeat. The interpretation of model 2 and 3 as in Table 4.4 is taken to avoid multicollinearity.

In Figure 4.3, the dashed horizontal line represents the effect of the total number interregional linkages. The dashed horizontal line is static because $LT_{r,t} = LC_{r,t} + LP_{r,t}$. The dashed vertical line represents the core-periphery dichotomy as in Table 4.4. The green line represents the regression coefficient for interregional linkages with core regions. The yellow line represents the regression coefficient for interregional linkages with peripheral regions. Interestingly, the effect of interregional linkages with core regions remains stable whilst adjusting the core-periphery dichotomy. However, the effect of interregional linkages with peripheral regions is highly variable. In particular, the effect of interregional linkages with peripheral regions shows a sigmoidal curve. At a split of 100% and beyond, the core and peripheral regions become comparable, and thus, from that split and on, the core-periphery dichotomy evens out.

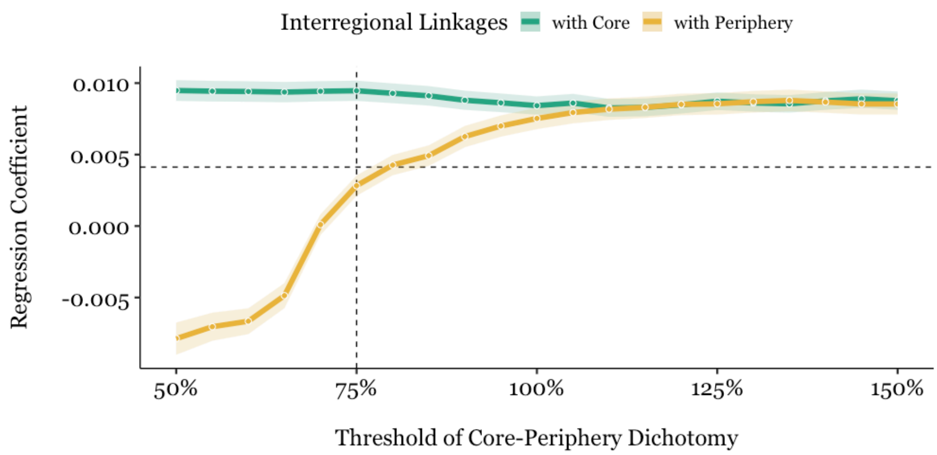


Figure 4.3: The parameterisation of the core-periphery dichotomy with the interpretation of model 2 and 3 as in Table 4.4. The shaded ribbon indicates three times the standard deviation around the regression coefficient.

In theoretical terms, core regions are stable providers of high quality external knowledge spillovers whilst peripheral regions are variable. Therefore, the far periphery would have less potential with regard to interregional collaboration. In contrast, core regions (no matter the split) do have potential to cause external knowledge spillovers. At the far periphery (below a split of 70%), the effect of interregional linkages with peripheral regions turns negative, which would suggest that linking to the far periphery obstructs technological diversification. Yet, another underlying cause could be that regions in the far periphery link together which represent a confounding effect of an inherent lock-in that is independent from measuring the actual effect of external knowledge spillovers from peripheral regions (i.e., a spurious relationship).

4.2.3. Stratification of the Core-Periphery Dichotomy

The next analysis stratifies core and peripheral regions in the central model. The stratification is of interest to investigate how interregional linkages affect technological diversification in core and peripheral regions differently. Table 4.5 indicates that interregional linkages with core and peripheral regions have different results for core regions from that in Table 4.4 because the effect of interregional linkages with peripheral regions loses significance in model 1 and 4. Therewith, the results in model 2 and 3 are similar to Table 4.4. In model 1, the effect of interregional linkages with peripheral regions does not seem to impact technological diversification in core regions. In model 4, the effect of interregional linkages with peripheral regions does not turn negative after adding the effect of interregional linkages with core regions. The results suggest that a difference exist in the flow of external knowledge spillovers as in Figure 4.4. All in all, a positive relationship exists (+) between the probability of a region to diversify and the number of interregional linkages with both core and peripheral regions, however, the effect of interregional linkages with peripheral regions has an insignificant effect (−) to core regions. Furthermore, the effect of interregional linkages with core regions is greater (++) to peripheral than to core regions.

In theoretical terms, whilst peripheral regions do benefit from interregional linkages with peripheral regions, core regions seem to not benefit from those. Thus, external knowledge spillovers seem to have a greater importance for peripheral regions with no difference to whom they link (may it be core or peripheral regions). Those results are crucial for policy-making in peripheral regions. If interregional linkages between core and peripheral regions does not benefit core regions, but solely benefits the peripheral regions, then the results would suggest the benefits of interregional linkages is unequal over the core-periphery dichotomy. From a positive perspective, peripheral regions benefit from external knowledge spillover regardless of whether the interregional linkages are with core and peripheral regions. Furthermore, no distinction is made between technological fields and external knowledge spillovers, and Figure 4.4 may revolve for technological field specific effects.

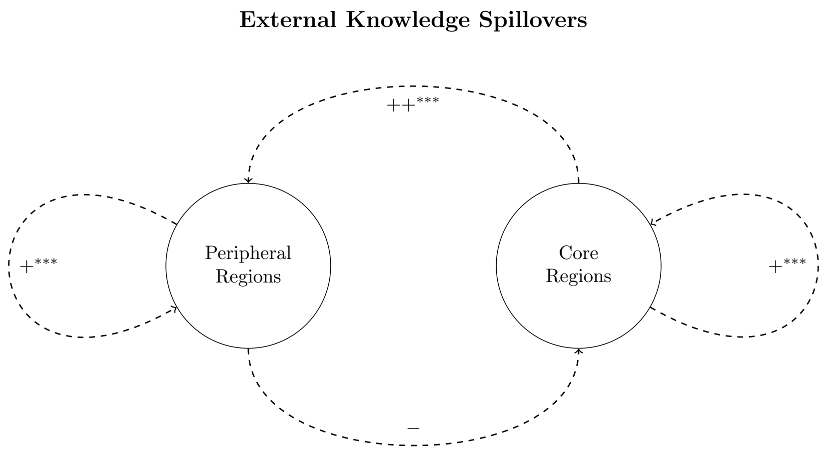


Figure 4.4: A visual summary of the interpretation of the effect of interregional linkages with core and peripheral regions for core and peripheral regions separately considering the sign of coefficients as in Table 4.5. The dotted lines represent the flow of external knowledge spillovers.

The insignificant effect of interregional linkages with peripheral regions to core regions can be interpreted as following. The overall effect of interregional linkages with peripheral regions to core regions is negligible. However, the effect of interregional linkages may vary over inherent characteristics of technological fields. In this thesis the interregional linkages are taken as a sum of all technological fields, and therefore no distinction is made between external knowledge spillovers for each technological field specifically. Further research could apply the specific definitions of interregional linkages Balland and Boschma (2021) for allowing the variance within technological fields. The exploratory analysis investigates further how collaboration between core and peripheral regions occurs over technological fields.

Table 4.5: The stratification of core and peripheral regions from the central model as in Table 4.4. All the models refer to Equation 3.4 (OLS) and include time fixed-effects α_t.

	Dependent variable:			
	$Entry_{i,r,t}$			
	Core		Periphery	
	(1)	(2)	(3)	(4)
Relatedness Density	0.004***	0.004***	0.003***	0.003***
	(0.00002)	(0.00002)	(0.00004)	(0.00004)
GDP per Capita	0.000***	0.000***	0.000***	−0.000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Population Size (log)	0.013***	0.008***	0.007***	0.001**
	(0.001)	(0.001)	(0.001)	(0.001)
Linkages Core (log)		0.007***		0.013***
		(0.0004)		(0.001)
Linkages Periphery (log)	−0.0004	−0.003***	0.010***	0.001
	(0.0003)	(0.0003)	(0.0004)	(0.001)
Period	Included	Included	Included	Included
Constant	−0.155***	−0.119***	−0.149***	−0.054***
	(0.008)	(0.008)	(0.008)	(0.009)
Observations	828,518	828,518	328,858	328,858
R^2	0.035	0.036	0.051	0.052
Residual Std. Error	0.337	0.337	0.254	0.254
F Statistic	1,789.296***	1,706.712***	1,031.806***	1,004.863***
Note:			*p<0.1; **p<0.05; ***p<0.01	

4.2.4. Model Evaluation & Diagnostics

The following paragraph contains a discussion in terms of the model evaluation & diagnostics. The model diagnostics revolves around specific tests on residuals and autocorrelation. These results on Breusch-Pagan tests, Durbin-Watson tests, and Moran tests can be found in the R code online at https://github.com/SebastiaanK97/EconomicGeography.

The evaluation of the central OLS model as in Table 4.4 is as following. Overall, the model fit by R^2 and the residual standard errors remain equal over model 1 to 7. Note, the model fit by R^2 is low, and the residual standard errors are high (~ 0.3) considering that the dependent variable range takes upon a value of either 0 or 1. In other words, the average misfit of the predicted value to the actual value is $\sim 30\%$ of the total range. Despite of that, the aim of the model is causal inference of the independent variables, and not achieving a high model fit. Regarding the model fit by F-statistics, model 2 (including the number of interregional linkages with core regions only) is a slightly better fit than model 1 (including the total number of interregional linkages only), and model 3 (including the number of interregional linkages with peripheral regions only) is the worst fit out of these three models. In line with previous results, the interregional linkages with core regions explains technological diversification best.

The diagnostics of the central OLS model as in Table 4.4 is as following. Analysis of the Breusch-Pagan test indicates that significant heteroskedasticity is present in all models (and also in robustness checks). The test indicates that the the model leads to unequal variance of residual error terms along independent variables. Likely, the definition of the dependent variable causes the violation of residuals to be constant in variance (Breusch & Pagan, 1980). Potentially, the independent variable is highly deterministic to assume that the entry of a technological field in a region is a binary variable. Nevertheless, the entry model is seen as a state-of-the-art model due to the ability to omit the bias in patent data well. Another issue that causes heteroskedasticity are spatial and temporal autocorrelation (Getis, 2007; Watson & Durbin, 1951). Analysis of the Durbin-Watson and Moran tests indicates that both significant spatial and temporal autocorrelation could be present in all models (and also in robustness checks). In other words, the entry of a technological field in one year or a neighbouring regions correlates with the entry of a technological field in the following year or the region itself. All in all, introducing additional variables or interactions may reduce autocorrelation.

Bear in mind, one should remain sceptic on the interpretation of regression coefficients. Specifically, how the effect of interregional linkages is attributable to further causes remains unknown i.e., are external knowledge spillovers measured by interregional linkages a cause of technological diversification, or, does a underlying pattern exist that confounds the effect of interregional linkages? For instance, one should question how the effect of interregional linkages is further attributable to different factors of knowledge production e.g., hard and soft infrastructure (Bathelt et al., 2004), or, the quality of government (Cortinovis et al., 2017). The discussion continues to list potential further causes and reflects on the bias in this thesis. Nevertheless, the robustness testing (subsection 4.2.5) of four further models indicate that the results withstand under various circumstances and alternative explanations.

4.2.5. Robustness Testing

The next step is performing robustness checks to investigate whether the hypotheses withstand different circumstances. As following, four robustness checks are performed. First, a robustness check is done by adding more variables and interaction effects to rule out various explanations. Second, a robustness check is done by adding 2-digit NUTS (country) and CPC (technological class titles) to adjust for country and technology specific differences. Third, a robustness check is done by stratifying the sample over time to regulate for fluctuations in time. Fourth, a robustness check is done by replacing ordinary linear regression with logistic regression to overcome inconsistencies in regression coefficients and to reduce the observational error of an entry (or not). Prior research by Balland and Boschma (2021) already shows the model is robust to zero-inflated negative binomials and adjusting the entry of a new technological field in a region from $RTA > 1$ to $RTA > 2$. Further, testing quadratic transformations are not relevant for the hypotheses since plenty literature suggests a linear positive effect of external knowledge spillovers on technological diversification.

The first robustness check takes two new variables and two interaction effects into account. First, the population density of a region is controlled for to rule out the effect of agglomeration economies. Namely, dense populations causes increased interactions between the population which in turn increase the probability of sharing knowledge between e.g., inventors (Frenken et al., 2007). This effect of interaction and sharing of knowledge is expected to have a positive relationship with technological diversification (Boschma et al., 2015). However, Table 4.6 indicates that the population density results into a significant negative effect.[5] Likely, the variance of population density to have a positive effect is taken over by other variables as a consequence of over-controlling which introduces multicollinearity into the models. Second, the R&D expenditure of a region is taken into account to control for a regions' inventive capacity. A region investing into technological advancements is expected to have increased technological diversification. Indeed, the variable yields a positive significant effect.[5] Third, the interaction effect between interregional linkages and the relatedness density is added to confirm whether interregional linkages may overcome the effect of relatedness density. In other words, the interaction effect represents whether external knowledge spillovers may over come the lock-in effects of path-dependency (Balland & Boschma, 2021).

Considering the two new variables and two interaction effects, the effect of both interregional linkages with core and peripheral regions remain significantly positive in model 1 and 2. Therefore, H1 remains true. Furthermore, the effect of interregional linkages with core regions also takes over the effect of interregional linkages with peripheral regions in model 3 and 4. Therefore, H2 remains true. All in all, interregional linkages overcome the confounding effects of agglomeration economies, inventive capacity, and its relation to path-dependency.

[5]The results remain stable considering all four models in Table 4.6.

In addition to the confounding effect by alternative explanations, model 3 and 4 in Table 4.6 compares the model with and without time fixed-effects i.e., the variable period. No significant differences are apparent apart that the interregional linkages with peripheral regions turns sign, nevertheless its effects in model 3 and 4 are insignificant.

Table 4.6: Robustness check by adding variables and interaction effects.

	Dependent variable:			
	$Entry_{i,r,t}$			
	(1)	(2)	(3)	(4)
Relatedness Density (RD)	0.002***	0.004***	0.002***	0.002***
	(0.0001)	(0.00005)	(0.0001)	(0.0001)
GDP per Capita	0.000***	0.000***	0.000**	0.000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Population Size (log)	0.002***	0.007***	0.003***	0.004***
	(0.0005)	(0.0005)	(0.0005)	(0.001)
Population Density (log)	−0.000***	−0.000***	−0.000***	−0.003***
	(0.00000)	(0.00000)	(0.00000)	(0.0003)
R&D Expenditure (log)	0.001***	0.003***	0.001***	0.001***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Linkages Core (log) - LC	0.006***		0.006***	0.005***
	(0.0003)		(0.0004)	(0.0004)
Linkages Periphery (log) - LP		0.002***	−0.001	0.001
		(0.0003)	(0.0005)	(0.0005)
RD:LC	0.0003***		0.0003***	0.0004***
	(0.00001)		(0.00001)	(0.00001)
RD:LP		0.00004***	−0.0001***	−0.0001***
		(0.00001)	(0.00002)	(0.00002)
Period	Included	Included	Included	Excluded
Constant	−0.036***	−0.118***	−0.055***	−0.048**
	(0.006)	(0.006)	(0.006)	(0.009)
Observations	1,113,635	1,113,635	1,113,635	1,113,635
R^2	0.046	0.044	0.046	0.046
Residual Std. Error	0.313	0.313	0.313	0.313
F Statistic	2,698.615***	2,592.229***	2,460.080***	6,004.812***
Note:				*p<0.1; **p<0.05; ***p<0.01

The second robustness check takes country and technological class fixed-effects into account. Here, controlling for countries is useful to attribute differences in country population their perception to the government (Cortinovis et al., 2017), and their attitude towards entrepreneurship and innovation (McCann & Ortega-Argilés, 2015). Additionally, controlling for technological classes is helpful to attribute differences in technological complexity i.e., tacit knowledge (Balland et al., 2019). Table 4.7 indicates that including the fixed-effect of 2-digit NUTS (countries) and CPC (technological class) leads to no major changes. Solely, the effect of economic development (GDP per capita) turns sign in model 1 and 3. Furthermore, the effect of interregional linkages with peripheral regions is positive in model 3, yet, remains smaller than the effect of interregional linkages with core regions. It should be mentioned that the estimation of 2-digit NUTS (countries) and CPC (technological class) are rough indicators due to the higher aggregation of regions and technological fields their 4-digit specification.[6]

Table 4.7: Robustness check by adding country ϕ_c and technological class ψ_i fixed-effects.

	Dependent variable:			
	$Entry_{i,r,t}$			
	(1)	(2)	(3)	(4)
Relatedness Density	0.004***	0.004***	0.004***	0.004***
	(0.00002)	(0.00002)	(0.00002)	(0.00002)
GDP per Capita	−0.000**	0.000***	−0.000**	0.000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Population Size	0.004***	0.011***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.0005)
Linkages Core (log)	0.009***		0.008***	0.010***
	(0.0004)		(0.0005)	(0.0003)
Linkages Periphery (log)		0.004***	0.001***	−0.002***
		(0.0003)	(0.0004)	(0.0003)
Period	Included	Included	Included	Excluded
Country	Included	Included	Included	Excluded
Technology	Included	Included	Included	Excluded
Constant	−0.065***	−0.123***	−0.056***	−0.071***
	(0.007)	(0.007)	(0.008)	(0.006)
Observations	1,157,376	1,157,376	1,157,376	1,157,376
R^2	0.049	0.049	0.049	0.045
Residual Std. Error	0.315	0.315	0.315	0.315
F Statistic	1,026.624***	1,021.208***	1,009.440***	10,982.970***
Note:				*p<0.1; **p<0.05; ***p<0.01

[6] Including fixed-effects at 4-digits of 176 regions and 621 technological fields is computationally impossible.

The third robustness check stratifies two periods of time from the central model. The stratification is of interest to investigate how variance in time affects technological diversification. Table 4.8 indicates that all the results are robust to splitting the time frame of $2005 - 2018$ over two equal sub-periods $2005 - 2011$ and $2012 - 2018$ because none of the variables change in sign. Arguably, the interregional linkages have become less important over time as the regression coefficients have become smaller. Potentially and underlying reason would be that regions have become better developed in e.g., hard and soft infrastructure, and thus the effect of interregional linkages may have become less apparent over time.

Table 4.8: Robustness check by stratifying two periods of time.

	Dependent variable:			
	$Entry_{i,r,t}$			
	2005-2011		2012-2018	
	(1)	(2)	(3)	(4)
Relatedness Density	0.004***	0.004***	0.004***	0.004***
	(0.00003)	(0.00003)	(0.00003)	(0.00003)
GDP per Capita	0.000***	0.000***	0.000***	0.000*
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Population Size (log)	0.011***	0.004***	0.010***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
Linkages Core (log)		0.013***		0.009***
		(0.0004)		(0.0004)
Linkages Periphery (log)	0.004***	−0.003***	0.001***	−0.002***
	(0.0003)	(0.0004)	(0.0004)	(0.0004)
Period	Included	Included	Included	Included
Constant	−0.164***	−0.093***	−0.126***	−0.058***
	(0.008)	(0.008)	(0.009)	(0.009)
Observations	581,011	581,011	576,365	576,365
R^2	0.046	0.047	0.043	0.044
Residual Std. Error	0.310	0.310	0.321	0.321
F Statistic	2,771.339***	2,617.129***	2,574.200***	2,394.871***
Note:			*p<0.1; **p<0.05; ***p<0.01	

The fourth robustness check replaces linear regression with logistic regression to test the hypotheses under different statistical assumptions (Li et al., 2020). Specifically, linear regression is not commonly applied when dealing with a binary dependent variable because logistic regression bounds the range between 0 and 1. Therefore, linear regression leads to biased regression coefficients and measurement errors in the dependent variable. However, logistic regression suffers from biased regression coefficients due to inconsistent model optimisation when dealing with large number of fixed-effects and too few degrees of freedom. In contrast, linear regression is not restricted in the usage of fixed-effects. Nevertheless, the overall effects of linear regression models are alike to logistic regression models. Therefore, linear regression is commonly applied in the literature of technological diversification (Balland et al., 2019; Boschma et al., 2015). Table 4.9 indicates the results logistic regression are similar to the results of linear regression as in Table 4.4 because none of the variables change in sign.

Table 4.9: Robustness check by logistic regression.

	Dependent variable:			
	$Entry_{i,r,t}$			
	(1)	(2)	(3)	(4)
Relatedness Density	0.030***	0.030***	0.032***	0.030***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
GDP per Capita	0.000***	0.000***	0.000***	0.000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Population Size (log)	0.026***	0.030***	0.149***	0.057***
	(0.005)	(0.005)	(0.005)	(0.005)
Linkages Total (log)	0.163***			−0.048
	(0.003)			(0.031)
Linkages Core (log)		0.160***		0.225***
		(0.003)		(0.028)
Linkages Periphery (log)			0.040***	−0.038***
			(0.002)	(0.004)
Period	Included	Included	Included	Included
Constant	−4.292***	−4.294***	−5.410***	−4.640***
	(0.060)	(0.060)	(0.062)	(0.065)
Observations	1,157,376	1,157,376	1,157,376	1,157,376
Log Likelihood	−394,765.500	−394,647.600	−396,466.000	−394,530.000
Akaike Inf. Crit.	789,567.000	789,331.200	792,968.000	789,100.000
Note:				*p<0.1; **p<0.05; ***p<0.01

4.3. Exploratory Analysis

As following is an exploratory analysis of both regions and technologies as separate units. The analysis aims to uncover how core and peripheral regions situate in terms of technological diversification and interregional linkages, and how interregional collaboration occurs.

4.3.1. Regional Analysis

To investigate regions further, the relative amount of entries and the relative GDP per capita are of interest to investigate the peripheral regions that overachieve in technological diversification. Therefore, aggregations are made over all regions over the entire period (the standard deviations that arise from aggregating time are low). The regions are classified on basis of quadrants that separate core and peripheral regions (75% of the average GDP per capita) and diversifying and non-diversifying regions (mean entries over all regions, which is 9.15%) to separate the regions into over- and underachievers according to their economic development. Figure 4.5 depicts the points of data within the four quadrants. Apparent is that peripheral regions are less likely to diversify than core regions, but also show the lowest amount of technological diversification. The quadrant of interest is that of diversifying peripheral regions (green) to understand how these regions situate. To take the analysis further the regions are mapped to their geography, and as a network of interregional linkages.

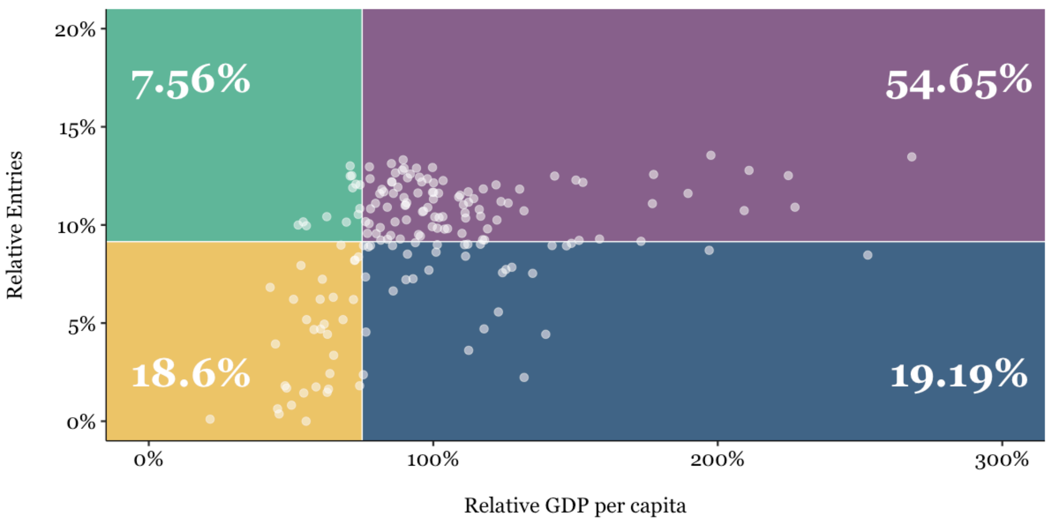


Figure 4.5: The representations of regions over the variables of economic development and technological diversification. The classification of the four quadrants clockwise are as following; purple depicts 94 diversifying core regions, blue depicts 33 non-diversifying core regions, orange 32 depicts non-diversifying peripheral regions, and green depicts 13 diversifying peripheral regions.

The diversifying peripheral regions are as following. Brandenburg, Chemnitz, Lüneburg, Mecklenburg-Vorpommern, Sachsen-Anhalt, and Thüringen in Germany. Andalucía and Comunidad Valenciana in Spain. Languedoc-Roussillon, Lorraine, and Picardy in France. Campania and Puglia in Italy. The diversifying peripheral regions are a compelling case for further qualitative analysis (e.g., comparative analysis or interviews) to understand how these regions overcome low knowledge production and do not restrict to a lock-in.

To create transparency in the data, Figure 4.6 visualises the quadrants geographically. In doing so, Figure 4.6 clarifies that the core and the peripheral regions cluster geographically. Especially the non-diversifying peripheral regions reside together and therewith add up to peripheral countries. The countries with the majority of non-diversifying peripheral regions are Belgium, Italy, Portugal, and Spain. The remaining peripheral regions reside in France and Germany. Likely, external knowledge spillovers, but also the quality of government, at the regional-level reflects back upon the country-level that causes technological diversification (Balland & Boschma, 2021; Cortinovis et al., 2017). In this thesis, clustering could be attributed to e.g., a lack of cultural, geographical, and institutional proximity (Boschma, 2005).

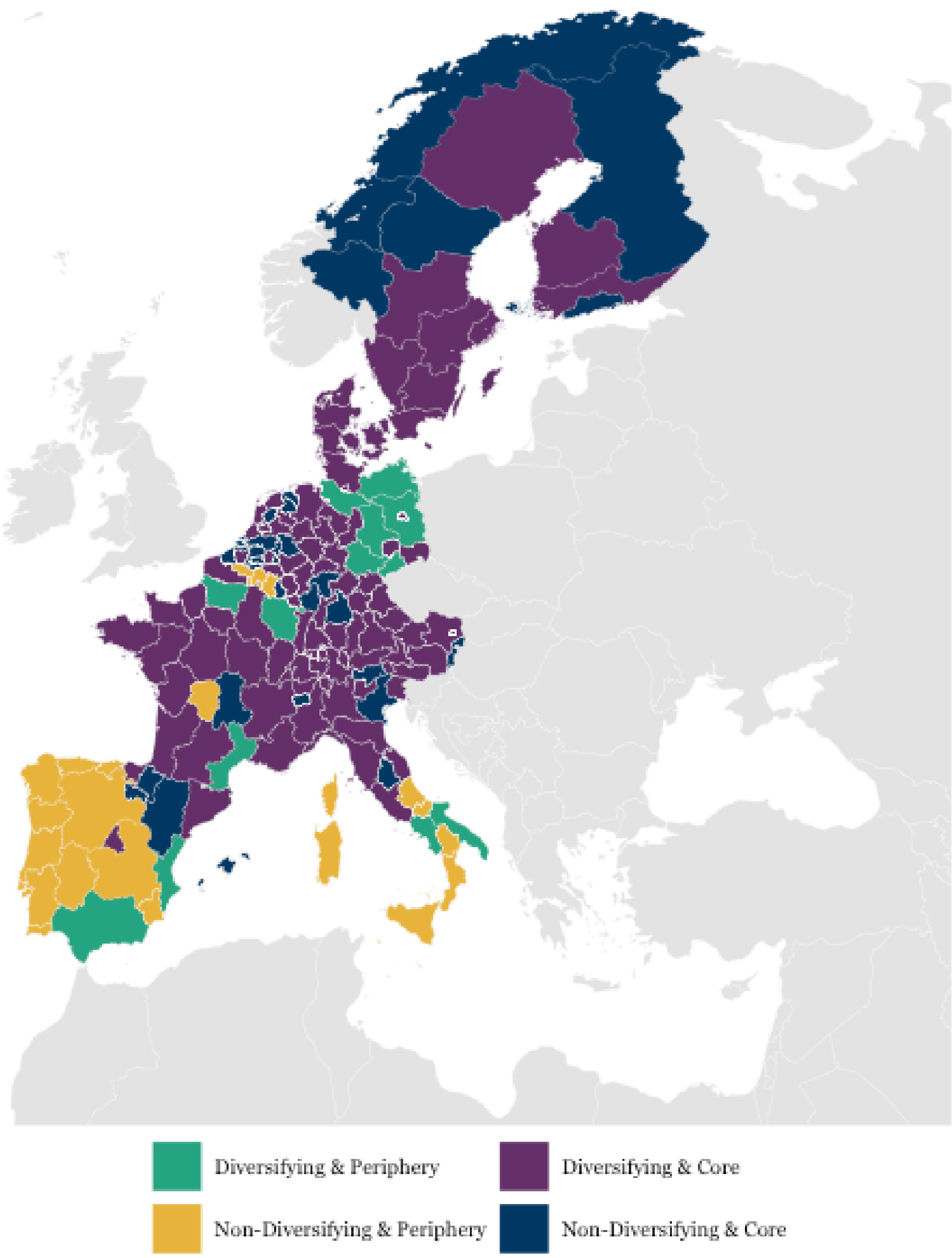


Figure 4.6: The geographical representations of the quadrants as in Figure 4.5.

Figure 4.7 depicts the regions as a network of co-invention. The network is set up using the quadrants to distinguish the regions (nodes) that collaborate through interregional linkages (edges). The SNA is useful to understand how regions draw upon external knowledge spillovers. Here, the network indicates that the core regions (by economic development) link closely together in the core of the network, further, the peripheral regions (by economic development) are also the periphery of the network. However, the peripheral regions that diversify situate less peripheral than those peripheral regions that do not diversify as much. Nevertheless, the explanatory results of Table 4.4 have already shown how interregional linkages affects technological diversification. The following exploratory analysis draws upon metrics of SNA to explore the importance of interregional linkages further.

Table 4.10 illustrates the mean of three centrality metrics of the SNA for each quadrant. First, in line with the explanatory results of Table 4.4, the degree centrality (i.e., the total number of interregional linkages) is higher for regions that diversify regardless of being a core or peripheral region. However, the non-diversifying core regions (blue) yield a much higher degree centrality than non-diversifying peripheral regions (orange). Therefore, the core regions benefit relatively more from external knowledge spillovers than the peripheral regions do. Second, the eigenvector centrality (i.e., the extent to which a region's neighbours connect) is exceptionally low for the non-diversifying peripheral regions (orange). This result indicates that the non-diversifying peripheral regions (orange) have low connectivity to other regions overall, and thus are the far periphery of the social network. Therefore, the exposure to a variety of external knowledge is lowest for non-diversifying peripheral regions (orange). This could be the exact reason why these regions tend to lock-in into their narrow knowledge base (Caragliu et al., 2016). Last but not least, the betweenness centrality (i.e., the shortest path to other regions) is remarkably large for the diversifying peripheral regions (green). Therefore, the diversifying peripheral regions (green) act as a bridge between regions. Further investigation of Table 4.10 clarifies that the diversifying peripheral regions intermediate between the core and the peripheral regions. In other words, external knowledge spillovers between the non-diversifying peripheral regions and all the core regions are indirect. One could say that diversifying peripheral regions are the gatekeepers of knowledge flow between the non-diversifying peripheral regions and all the core regions. Why diversifying peripheral regions are gatekeepers remains unknown. Further research could focus on how to attribute the variance in betweenness centrality (as a form of power relations) to factors such as entrepreneurship, labour markets (and migration), or flows of natural resources (Nilsen et al., 2023).

Table 4.10: Mean centrality metrics of the network in Table 4.10 per quadrant.

Quadrant	Entries	Degree	Eigenvector	Betweenness
(1) Diversifying & Core	11.17%	110	0.32	68
(2) Non-Diversifying & Core	7.42%	83	0.23	66
(3) Non-Diversifying & Periphery	4.16%	40	0.04	66
(4) Diversifying & Periphery	11.24%	98	0.25	84

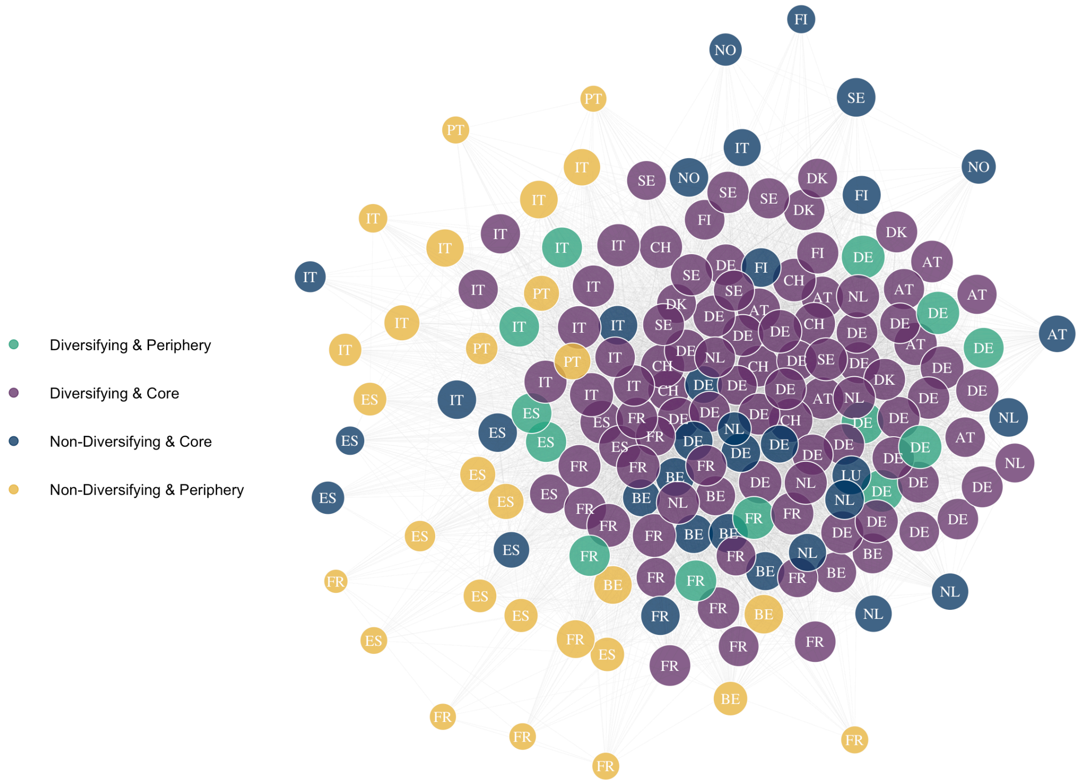


Figure 4.7: Network depicting the regions as nodes and interregional linkages as edges with an identical indication of the quadrants as in Figure 4.5.

Figure 4.8 depicts the assortativity coefficient over time to understand how core and peripheral regions draw upon external knowledge spillovers.[7] Here, interregional collaboration between core and peripheral Western European regions is (slightly) homophilous. Thus, core regions tend to draw upon external knowledge spillovers from other core regions whilst peripheral regions tend to draw upon external knowledge spillovers from other peripheral regions. Furthermore, Figure 4.8 indicates that interregional linkages become more homophilous over time. Further investigation signifies that especially core regions increasingly link to other core regions whilst the development of interregional linkages in peripheral regions remains rather stable over time. In other words, core regions increase their access to high quality external knowledge spillovers whilst peripheral regions are rigid and do not.

A straightforward explanation would be that core regions reside in core countries. The core regions in core countries increase in their interregional collaboration whilst peripheral regions remain rigid. The general conception would be that peripheral regions face excessive barriers to reach out to an external network of co-invention due to e.g., cultural and geographical, institutional, and organisational proximity (Boschma, 2005). How these barriers could be attributed further to factors of e.g., hard and soft infrastructure (Bathelt et al., 2004), or the quality of government (Cortinovis et al., 2017) would be appealing for further research.

From the perspective of economic equality, policy should intervene and aim for a heterophilous network of interregional collaboration. Especially due to the results in Table 4.4 indicate that interregional linkages with core regions cause high quality external knowledge spillovers in comparison to interregional linkages with peripheral regions. The peripheral regions should become able to draw upon external knowledge spillovers from core regions to restore an equilibrium in external knowledge spillovers and thus technological diversification.

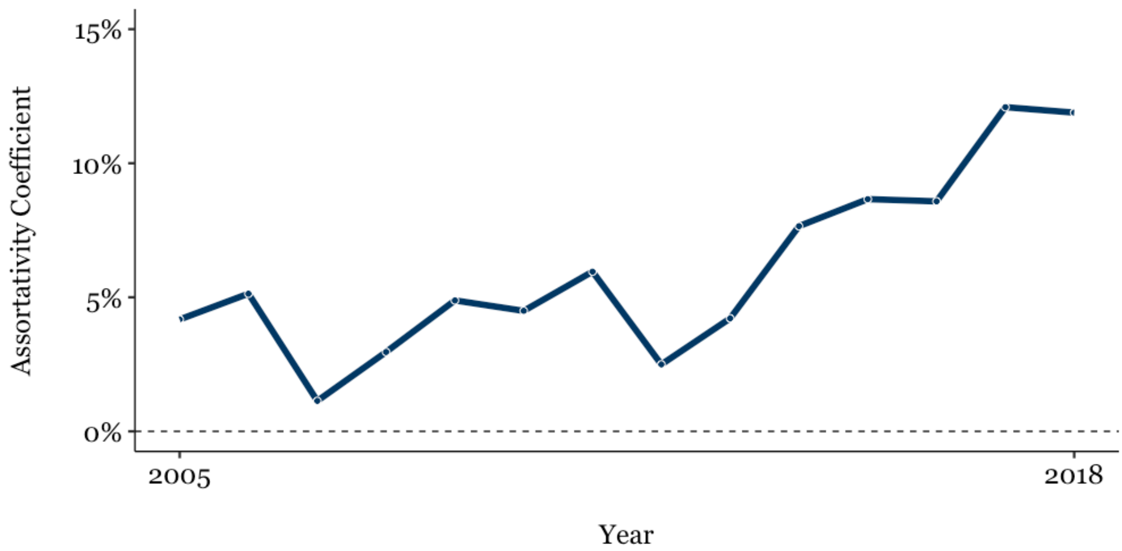


Figure 4.8: Assortativity coefficient of the core-periphery dichotomy over time.

[7] Whereas 0% would indicate a random network, −100% indicates heterophily (i.e., core regions only link to peripheral regions, and vice versa, peripheral regions only link to core regions), and +100% indicates homophily (i.e., core regions only link to core regions, and vice versa, peripheral regions only link to peripheral regions)

4.3.2. Technological Analysis

To investigate technological fields further, the difference in entries as in Table D.1 is taken in conjunction with the assortativity coefficient specific for each technological field over the entire period of $2005 - 2018$. The technological fields are classified on basis of quadrants that separate homophilous and heterophilous technological fields and technological fields that are dominant in either core or peripheral regions. Figure 4.9 depicts the points of data within the four quadrants. Apparent is that the majority of the technological fields is dominant in core regions and the majority of technological fields is homophilous. In other words, the majority of technological fields arise in core regions through collaboration between core regions. The quadrant of interest is the technological fields that arise dominantly in peripheral regions through heterophilous collaboration (thus with core regions) because the results in Table 4.4 indicate that such collaboration induces technological diversification. To take the analysis further the technological fields are mapped in the technology space and metrics are analysed.

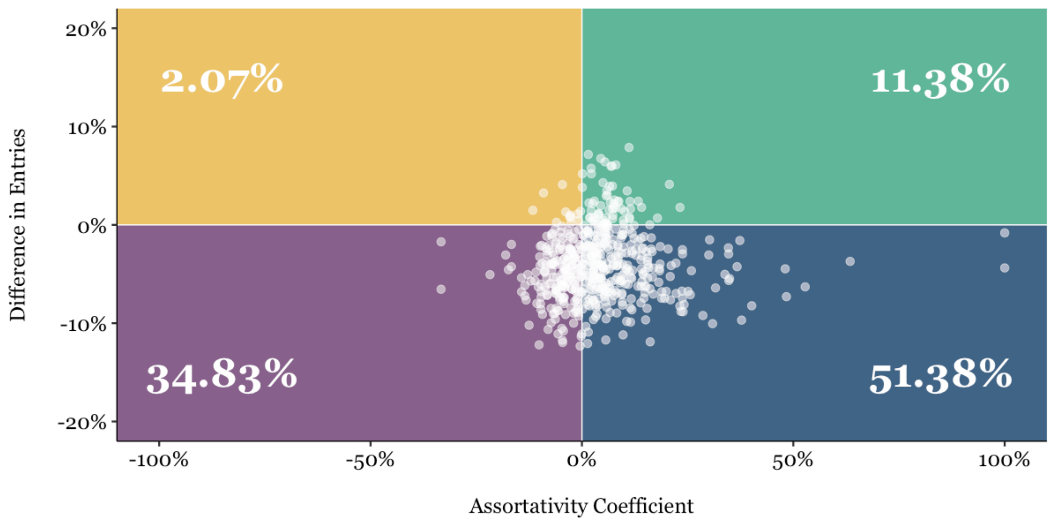


Figure 4.9: The representations of technological fields over the variables of assortativity and core-periphery dominance. The classification of the four quadrants clockwise are as following; green depicts 66 homophilous peripheral technological fields, blue depicts 298 homophilous core technological fields, purple 202 depicts heterophilous core technological fields, and orange depicts 12 heterophilous peripheral technological fields.

The heterophilous peripheral dominant technological fields are as following. Dentistry; Oral Dental Hygiene (A61C) in Human Necessities. Vehicle Suspension (B60G) and Motor Vehicles; Trailers (B62D) in Performing Operations & Transporting. Macromolecular Compounds; Carbon-to-Carbon (C08F), Cracking Hydrocarbon Oils (C10G), Detergent Compositions (C11D), and Manufacture of Iron Steel (C21B) in Chemistry & Metallurgy. Paper-Making Machines (D21F) in Textiles & Paper. Locks; Accessories; Handcuffs (E05B) and Earth Drilling (E21B) in Fixed Constructions. Ignition (F02P) and Domestic Stoves or Ranges (F24B) in Mechanical Engineering. These technological fields are a compelling case for further qualitative analysis (e.g., comparative analysis or interviews) to understand why interregional collaboration occurs between these technological fields whilst benefiting peripheral regions.

As Hassink and Gong (2019) state "[peripheral] regions should also have policy portfolios to support their horizontal research and innovation capabilities, so that they would not run the risk of further lagging behind in the current and future rounds of the digital and knowledge economy." Therefore, peripheral regions should diversify into any technological field possible (regardless of complexity, novelty, or ubiquity). Thus, considering the external knowledge spillovers from core regions (as proven in Table 4.4), policy should support heterophilous technological fields to enable knowledge spillovers across the core-periphery dichotomy.

To understand what general contrasts exist, Figure 4.10 indicates the percentage to what the quadrants belong to a technological section. Apparent is that Chemistry & Metallurgy (C) and Fixed Constructions (E) are most dominant in peripheral regions. Probably, these technological sections reflect upon labour and natural resource intensive industries (Eder, 2019; Eder & Trippl, 2019). Further, homophilous peripheral regions are not presented in the technological section of Textiles & Paper (D). Besides, heterophilous peripheral regions are not presented in the technological sections of Physics (G) and Electricity (H). The under presentation in these technological sections is likely in line with the reasoning that solely core regions diversify into complex technological fields (Balland et al., 2020). No contrasts are found between homophilous and heterophilous collaboration in Figure 4.10.

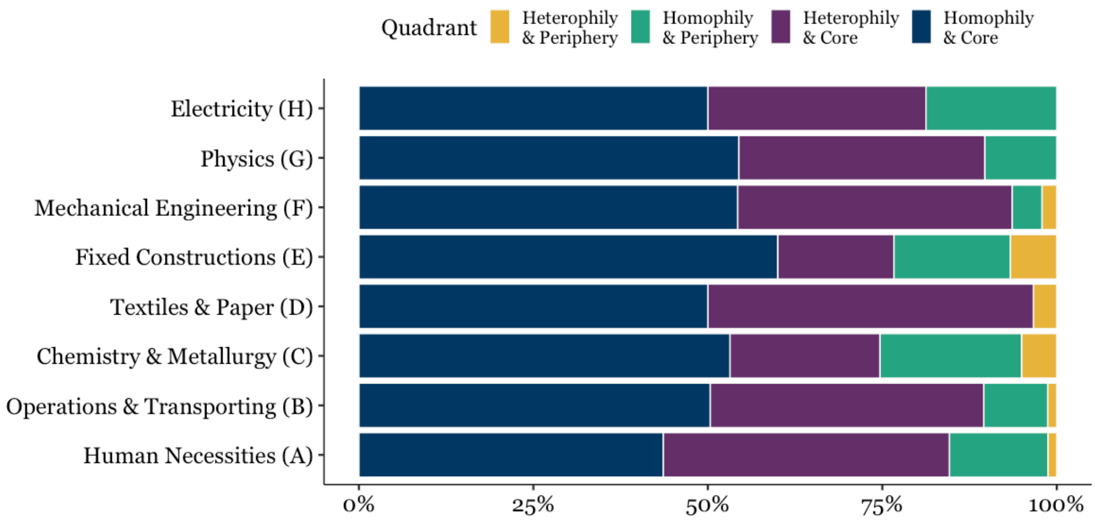


Figure 4.10: Presence of technological section per quadrant as in Figure 4.9.

Figure 4.11 depicts the technology space using the quadrants to distinguish technological fields (nodes) that tie together according to their relatedness (edges). The technology space is useful to depict how technological fields relate in a general sense. Here, the technology space indicates that the technological fields that relatively arise more in peripheral regions (orange & green) tend to cluster slightly. Further, the heterophilous peripheral dominant technological fields (orange) scatter randomly in the network. The periphery of the technology space i.e., novel technological fields, consists only of technological fields that arise dominantly in core regions (blue & purple). Figure 4.12 describes these findings further along theoretical terms.

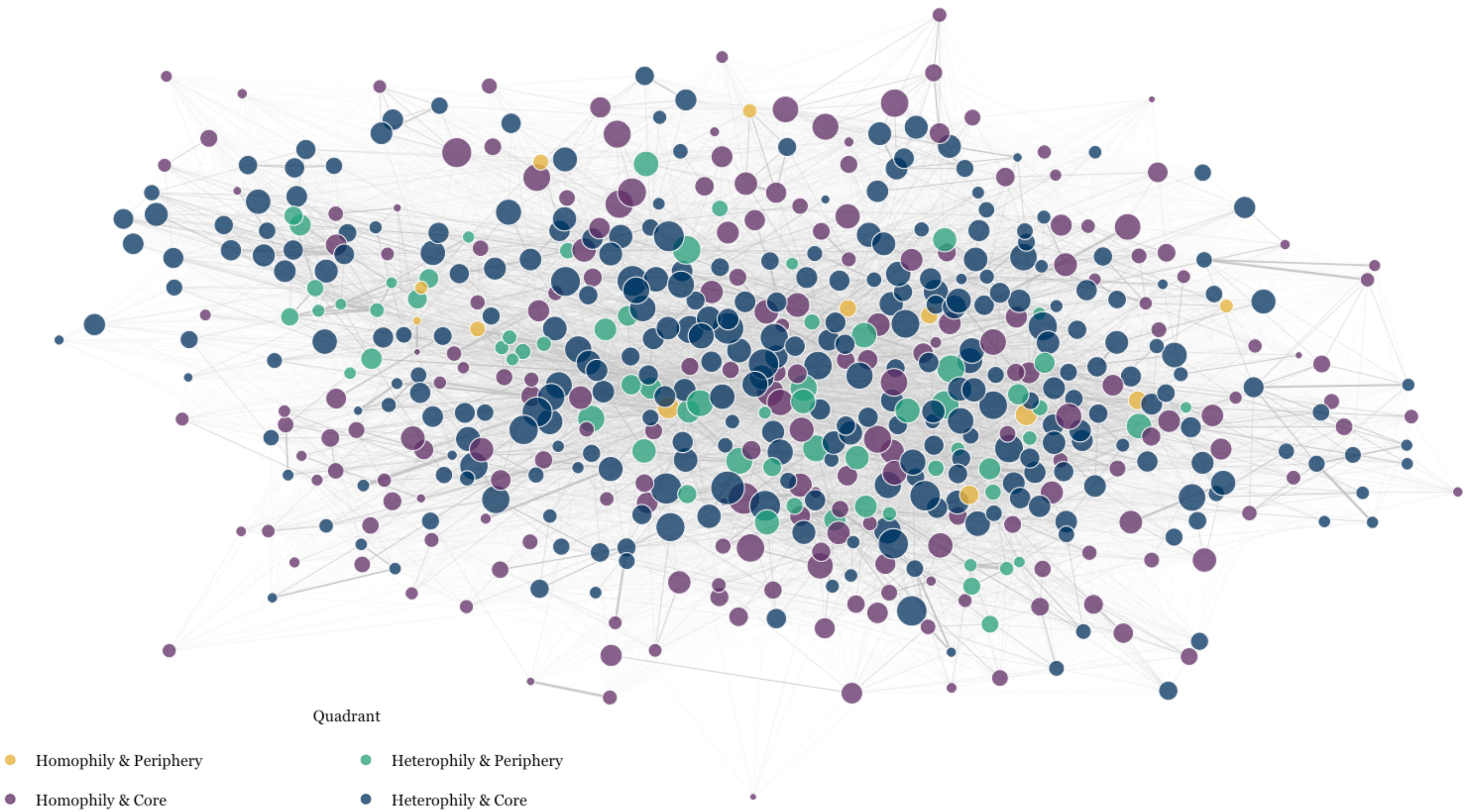


Figure 4.11: Technology space depicting the technological fields as nodes and relatedness as edges with an identical indication of the quadrants as in Figure 4.9.

Figure 4.12 depicts the mean novelty of technological fields per quadrant which is taken by the mean relatedness to other technological fields (Balland & Rigby, 2017; Fleming & Sorenson, 2001). Therewith, lower mean relatedness refers to higher novelty in regions. Figure 4.12 highlights two findings in a general sense; the technological fields that arise dominantly in peripheral regions are less novel, and the technological fields that arise through heterophilous collaboration are more novel i.e., those with core regions. These two findings are as follows:

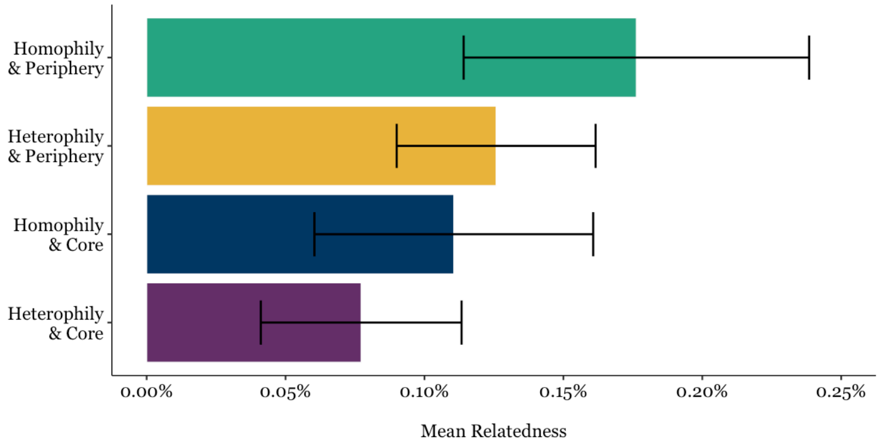


Figure 4.12: The novelty of technological fields per quadrant measured by the mean relatedness with the whiskers that represent the standard deviation around the mean.

First, peripheral regions face the challenge to catch-up to overall technological development in the global environment. On one hand, the major technological change in peripheral regions lies in modernisation towards general-purpose technologies to upgrade their economy at a basic level (Hassink & Gong, 2019). On the other hand, core regions already control such general-purpose technologies and upgrade their economy further towards novel and complex technological fields at the boundary of the technological frontier (McCann & Ortega-Argilés, 2015). All in all, core regions are able to diversify into novel technological fields whilst peripheral regions lag and therefore tend to diversify into general-purpose technologies. Notable is that the diversification into novel technological fields is risky and uncertain but also brings new market opportunities (Balland & Rigby, 2017). By definition, the core regions have benefits to deal with risk and uncertainty due to the available resources in those core regions.

Second, considering the prior results, one could expect that the introduction of novel technological fields would occur between core regions. However, Figure 4.12 highlights that the technological fields that arise through heterophilous collaboration are lower in their relatedness to other technological fields i.e., more novel. In other words, collaboration between core and peripheral regions leads to the introduction of novel technological fields. These findings are unexpected and unknown in theory of technological diversification. An explanation could be that the core and peripheral regions diverge in their knowledge bases and that the recombination of those knowledge bases leads to the creation of novel knowledge (Fleming & Sorenson, 2001). Another explanation could be that collaboration between core and peripheral regions leads to effective use of labour and natural resource intensive industries which enables the production of novel knowledge (Eder, 2019; Eder & Trippl, 2019).

Figure 4.13 depicts the mean ubiquity of technological fields per quadrant. The ubiquity is taken by the the relative entries of technological fields over all regions. Therewith, high relative number of entries in regions refers to high ubiquity whereas a low relative number of entries in regions refers to low ubiquity. In other words, a high ubiquity means that the entry of technological fields is available to many regions whereas low ubiquity means the entry of technological fields is rare in comparison to the sample. No notable differences are found in the technological ubiquity between the four quadrants.

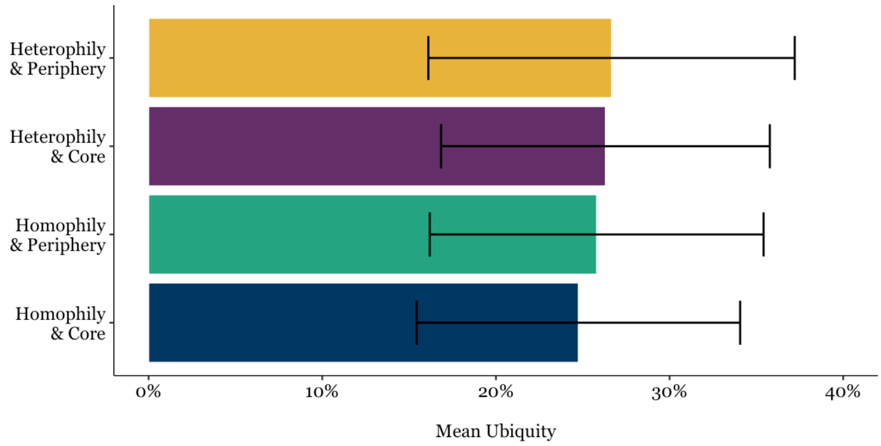


Figure 4.13: The ubiquity of technological fields per quadrant measured by the relative entries of regions with the whiskers that represent the standard deviation around the mean.

The findings would indicate that the technological fields within those quadrants are neither different in rarity nor availability. Therefore, the technological fields regions specialise in are just as present in core regions as in peripheral regions considering the mean ubiquity. Furthermore, interregional collaboration over the core-periphery dichotomy does not seem to cause a difference in the mean ubiquity of technological fields.

In summary, the regression models in thesis have shown to what extent interregional linkages affect technological diversification in core and peripheral regions differently. Further analysis shows how core and peripheral regions situate in a network of co-invention as of interregional linkages, and that interregional collaboration may lead to novel knowledge in the economy. The next chapter is the conclusion that recapitulates on theses findings.

5
Conclusion

This thesis builds upon the theoretical foundations of Evolutionary Economic Geography (EEG) literature particularly on how external knowledge spillovers enable technological diversification in especially peripheral European regions. Therewith, the main research question is:

How do the interregional linkages with core and peripheral regions affect the probability of a region to technologically diversify?

To address this research question, a quantitative explanatory and exploratory approach was taken to investigate technological diversification measured by the entry of a technological field (CPC4) in regions (NUTS2) between $2005-2018$. Herewith, external knowledge spillovers traced by interregional linkages were put central because of its presumed positive impact to peripheral regions. The explanatory analysis aimed to explain the importance of interregional linkages with core and with peripheral regions on technological diversification in regions. The exploratory analysis aimed to explore how technological diversification comes about in peripheral regions, and what technological fields show potential for technological diversification in peripheral regions. The results of both types of analyses revealed:

First, considering the regression models, it seems that interregional linkages with core regions are more beneficial to technological diversification than interregional linkages with peripheral regions. One possible explanation for this pattern could be that the cost of collaborating with peripheral regions is higher than collaborating with core regions because internal knowledge production is lower in peripheral regions. Another possible explanation could be that interregional linkages actually measures an underlying pattern of innovation that is partially independent from external knowledge spillovers because the interregional linkages of a region may relate to internal knowledge production itself. Additionally, analysis of the assortativity coefficient identifies that the number of interregional linkages between core regions has risen whilst peripheral regions are more rigid and have not increased their relative number of interregional linkages with core regions. Therefore, at this rate, external knowledge spillovers do not seem to close the gap in knowledge production between leading core and lagging peripheral regions. All in all, it is likely that core regions are less willing to collaborate with peripheral regions than with core regions due to the higher benefits of collaborating with those core regions. Thus, the importance of interregional linkages to peripheral regions is promising in theory but might not entirely enable peripheral regions to catch-up to core regions. Nevertheless, as shown by the regression models, external knowledge spillovers can reduce the tendency of peripheral regions to lock-in due to its positive effect on technological diversification. Therefore, interregional collaboration remains crucial for peripheral regions.

Second, the exploratory analysis of regions aims to identify patterns in the geography of regions and as network of co-invention. The most interesting findings concern peripheral regions that diversify since these regions are overachieving in terms of technological diversification in comparison to what is expected of their level of economic development. Therewith, the SNA of co-invention indicates that a notable difference exist between core and peripheral regions. In particular, the regions that are peripheral by their economic development are also the periphery in a network of co-invention. However, an interesting finding is that the betweenness centrality is remarkably large for peripheral regions that diversify. Likely, these regions function as a bridge between core and peripheral regions due to, for instance, speaking multiple languages, low travel distances, and an agreement between institutions. Herewith, these regions may have the benefit to access different labour markets (and migration), or flows of natural resources (Nilsen et al., 2023). A question for further research is how peripheral regions can close their gap in collaboration with core regions, and thus which type of proximity is most important to bring core and peripheral regions together (Boschma, 2005).

Third, the exploratory analysis of technological fields investigates patterns in collaboration between core and peripheral regions. It becomes apparent that core and peripheral regions have different benefits to different technological sections which is in line with that peripheral regions specialise into less complex technological fields (Balland et al., 2020). Further, the technology space among which the mean relatedness of technological fields indicates an interesting finding on how novelty arises in technological diversification. The findings suggest that heterophilous collaboration (i.e., interregional collaboration between core and peripheral regions) leads to diversification into novel technological fields in comparison to homophilous collaboration. These findings may relate to the understanding that distant (unrelated) knowledge recombination leads to novelty (Fleming & Sorenson, 2001) All in all, collaboration between core and peripheral regions could lead to the creation of novel knowledge that benefits society at the whole. However, these findings should be further tested in future research by performing explanatory analysis at the level of technological fields.

Overall, a scientific and societal debate is not settled on what the role is of policy in place-based bottom-up approaches (Iammarino et al., 2019). Another debate has mainly drawn attention to economic activity of firms and sectors, but nevertheless could also apply to regions. The statement by Mazzucato (2018) ”from picking winners to picking the willing” would translate in the context of this thesis into that the winners refer to core regions and the willing refers to any region where change can be made. Throughout history policy has been focusing on putting large organisations within core regions together because there technological change has been most prominent. This thesis promotes that the role of policy should shift from a preference to make technological change happen in core regions to regions where change in innovation can be made. Therewith, if interregional collaboration between core and peripheral truly leads to novel technological change then the role of policy could highlight peripheral regions better. Such policy would be beneficial for sustainable and inclusive growth across core and peripheral regions.

6
Discussion

As following is the discussion which consists out of four sections. First, the discussion starts with describing the theoretical contributions in terms of alignment to previous literature and recommendations for further research. Subsequently, the theoretical and the methodological limitations are explained. Thereafter, the methodological suggestions are described in terms of how the data and the entry model can be improved for future analysis. Lastly, the policy implications are drawn in the context of the EU Cohesion Policy and S3.

6.1. Theoretical Contributions

This thesis is contributing to the understanding of technological diversification in peripheral regions through explanatory and exploratory analysis. The explanatory analysis extends the literature on the effect of interregional linkages (i.e., external knowledge spillovers) on technological diversification (Balland & Boschma, 2021) by separating the effect of interregional linkages from those with core and peripheral regions. The exploratory analysis identifies to what extent core and peripheral regions make use of interregional linkages over the core-periphery dichotomy. Further, the exploratory analysis scrutinises interregional collaboration between core and peripheral regions in terms of the technology space.

Prior research has argued that interregional linkages enable technological diversification in regions. However, prior research has neglected to whom these regions link. The explanatory analysis indicates that interregional linkages with core and peripheral regions are not of equal importance due to the difference in knowledge production in those core and peripheral regions. In other words, a hierarchy may exist between low and high quality external knowledge spillovers. Thus, next to the argument of Balland and Boschma (2021) that interregional linkages are especially important to peripheral regions to overcome a lock-in into their existing knowledge base, this thesis argues that the level of economic development of the linking region is important as well for technological diversification. Therefore, in economic terms, costs are related to the economic development of interregional collaboration i.e., the cost of collaborating with core regions is lower than collaborating with peripheral regions due to factors of internal knowledge production in regions. These factors refer to the stylised facts such as the population density, the market size, the number of firms, the number of start-ups, and the number of multinational firms in a region. This theoretical contribution lead directly to policy implications on how to make interregional collaboration more inclusive over the core-periphery dichotomy.

Additionally, this thesis attempts to identify underlying patterns of how technological diversification is possible in peripheral regions (Eder, 2019; Eder & Trippl, 2019). The findings identify that peripheral regions that do diversify are exceptionally high in betweenness degree in comparison to core regions. The findings indicate that the peripheral regions that diversify situate between peripheral regions that do not diversify and all core regions. In other words, these regions could be seen as gatekeepers of knowledge between the knowledge production within core and peripheral regions. Further research could focus on how this relates to labour markets (and migration), or flows of natural resources (Nilsen et al., 2023). Therefore, further research could take the betweenness centrality of regions as the dependent variable with indicators on labour costs and natural resource intensity as the independent variables.

Further findings highlight that cultural, geographical, and institutional proximity may affect how interregional linkages occur (Boschma, 2005). Therewith, the importance of interregional linkages with core regions over those with peripheral regions brings us to the question on how to bridge the gap in proximity between core and peripheral regions, for example, in terms of language barriers (cultural proximity), long travelling distance (geographical proximity), and political or social differences (institutional proximity). Thus far, it remains unknown which proximity is most influential to link (European) core and peripheral regions (Balland & Boschma, 2021). Further research into proximity would directly relate to the policy implication of how to justify allocation of funding in improving, for instance, hard (e.g., highways and airports) and soft (e.g., high-speed internet) infrastructure, linguistic education, and populations their perceptions and experiences with the government. Therewith, it would be crucial to stratify the effects over core and peripheral regions when performing further quantitative explanatory research because these effects could differ over the core-periphery dichotomy.

Next to the level of regions this thesis has shown that at the level of technological fields differences exist that relate to collaboration between core and peripheral regions. In particular, the novel technological fields seem to arise in core regions which puts a negative emphasises on peripheral regions further. However, it seems that within the technological fields where high interregional collaboration occurs between core and peripheral regions novelty arises. In these terms of novelty, peripheral regions may lead to the production of novel knowledge. How these technological fields relate to e.g., labour and resource intensity would be appealing for further research (Eder, 2019; Eder & Trippl, 2019). A straightforward approach would be to take the novelty of technological fields as the dependent variable whereas the assortativity coefficient (considering core and peripheral regions as nodes and co-invention between regions as edges) as the independent variable. Therewith, controlling for the complexity of technological fields would be crucial to assign the statistical differences correctly (Balland et al., 2020).

6.2. Limitations

This thesis has several theoretical and methodological limitations. The theoretical limitations hold in the delimitation to technological change, and thus neglecting topics such as social innovation and political importance. The methodological limitations hold in the generalisation beyond the sample, the definition of interregional linkages, and further bias in patent data.

The theoretical limitations are as following. First, an important notice is that social innovation may be of greater importance than technological innovation in peripheral regions (Eder, 2019). Therefore, the emphasis on technological innovation shifts the centre of attention in this thesis towards a type of innovation that is difficult to achieve in peripheral regions. This limitation also reflects upon the methodological issues of taking patent data to measure innovation in peripheral regions (which is addressed further in this section).

Second, in relation to internal reliability, this thesis neglects the confounding effects which are here taken by the fixed-effects of countries and technological classes. The underlying pattern of how technological diversification relates to countries and regions could be measured by the quality of government (Cortinovis et al., 2017). However, statistical modelling on basis of this data is difficult since it is only measured in the years 2010, 2013, 2017, and 2021 (from which only three periods overlap the available patent data). Moreover, the underlying pattern of how technological diversification relates to technological classes and fields could be measured by technological complexity (Balland et al., 2019). Therewith, complex technological fields are often referred to tacit knowledge which is difficult to produce and thus could bring greater economic benefits. Herewith, the assumption is that new specialisations into complex technological fields is easier for core regions than for peripheral regions because core regions have a broader and deeper knowledge base. This thesis did not consider the complexity of technological field as an additional factor in the study of diversification which could be a limitation and issue for further research. However, the issue of complex technological fields need to be treated with caution in the context of peripheral regions. As argued by Hassink and Gong (2019), peripheral regions should prioritise specialising into any new technological field to broaden their knowledge base in general ”so that they would not run the risk of further lagging behind in the current and future rounds of the digital and knowledge economy” (p. 8).

Third, this thesis investigates the core-periphery dichotomy as a divide in poor and rich regions because GDP per capita is a solid and widely available indicator of knowledge production (Iammarino et al., 2019). However, different definitions of the core-periphery dichotomy in knowledge production and innovation could be thought off. As the six stylised facts indicate, other definitions could revolve around the population density, the market size, the number of firms, the number of start-ups, or the number of multinational firms in a region. According to Eder (2019), the population density would be another common measurement of the core-periphery dichotomy (as urban and rural regions besides rich and poor regions) to describe knowledge production and innovation. Furthermore, one could think of that, besides a general core-periphery dichotomy, also a core-periphery dichotomy would exist within nations (or other non-geographical boundaries). In other words, a region may be core within the general sample, but still be perceived as a peripheral region within different boundaries.

The methodological limitations are as following. First, in relation to external validity, the findings are delimited to the sample of Western European Regions, and worldwide, this sample is at the extremity of core regions in First World countries. How the explanatory findings (i.e., effects of interregional linkages) withstand under another spectrum of core and peripheral regions would be interesting in itself, and particularly to understand the costs associated with external knowledge spillovers better. But note, the eligibility of regions using patent data (i.e., a threshold of 50 patents per region per year) is bound to the level of economic development, and therefore researching technological diversification in regions with lower economic development would be controversial. Therefore, researching more peripheral regions would rely on different methodologies (e.g., surveys) or different theories (e.g., social innovation).

Second, due to the delimitation of the sample to Western European regions, also the interregional linkages delimit to those regions. In other words, co-invention does not exist outside Western European regions due to the data availability (of GDP per capita) to regions in the EU. Herewith, this thesis underestimates measuring external knowledge spillovers for certain regions that are deeply embedded into global value chains outside of Western Europe. On top of that, a small remark can be made that interregional linkages should not be counted for core and peripheral regions separately but should be measured as weighted according to the level of economic development of linking regions. The binary dichotomy eases theorisation and interpretations, yet, leads to biased and less significant results.

Third, in relation to internal validity, the entry model is a relative approach that makes regions comparable regardless of their number of patents. However, the entry model also overestimates the entry of a new specialisation in regions with very few patents. Even though a threshold of 50 patents per region per year is taken into account, the methodology remains sensitive to (peripheral) regions at the lower boundary of only 50 patents in a year (considering that 621 technological fields apply). Furthermore, the entry model neglects the depth of a specialisation in technological fields. Even though S3 emphasises that peripheral regions should diversify as of horizontal development, and not specialise as of vertical development, the relation between horizontal and vertical development could be of importance to technological change. Additionally, the entry model in this thesis does not consider the sustainability of a specialisation in a technological field in two ways. First, this thesis does not consider whether a specialisation in a technological field is sustainable regarding that these specialisations remain over multiple periods. Second, this thesis neither considers whether a specialisation in a technological field is sustainable in terms of mitigation or adaptation against climate change.

Fourth, a last remark should be made that collaboration and exploitation is not considered in measuring interregional linkages. This distinction refers to the dependency between core and peripheral regions. Thus, core regions may collaborate with peripheral regions meanwhile they exploit peripheral regions for e.g., their (low-cost) resources or labour. Therewith, in spite of that collaboration may lead to economic benefits to peripheral regions, collaboration may also have an adverse impact on society due to exploitation as of economic dependencies.

6.3. Policy Implications

This thesis is contributing to policy-making in the context of S3 in the EU. The S3 is a place-based approach that sets technology-driven priorities to solve socio-economic challenges. In these means, S3 considers that policy must be kept bottom-up and that technological change should allow for regional path-dependent development. Therewith, this thesis seeks to justify how the allocation of funding fosters external knowledge spillovers between core and peripheral regions by pinpointing different aspects of technological fields.

As stated before in this thesis, the lesser developed (i.e., peripheral) European regions require to diversify in order to induce economic competitiveness, expand the labour market, and transition towards a sustainable and digital economy (Balland et al., 2019). To enable technological diversification, this thesis emphasises that collaborating with core regions is of greater importance than collaborating with peripheral regions. Therewith, especially the peripheral regions benefit from collaborating with core regions. However, in relation to e.g., cultural, geographical, and institutional proximity, peripheral regions tend to collaborate with other peripheral regions whilst core regions dominantly collaborate reciprocally. To foster sustainable and inclusive growth across regions, the EU Cohesion Policy has a task to link core and peripheral regions together to not let the core-periphery dichotomy grow further.

This thesis identifies that several peripheral regions do diversify whilst many peripheral regions do not. The peripheral regions that do diversify are self-sufficient. Especially, the peripheral regions that do not diversify require aid from the EU Cohesion Policy. A generic solution would be that EU funding programmes target collaboration between peripheral regions that do not diversify and the core regions that provide complementary knowledge to each other (Balland & Boschma, 2021). But, how to make interregional collaboration between core and peripheral regions practical and effective, and, how to make interregional collaboration between core and peripheral regions attractive for core regions is a challenging task. A specific solution could be to target collaboration between core and peripheral regions whereas the peripheral regions also cause a positive contribution (i.e, external knowledge spillovers) to core regions. These solutions would be specific to technological fields and requires further consideration their particular circumstances.

The overlapping issue is that policy-makers in the EU also have the responsibility to reduce the productivity gap between the EU to other nations from which especially the United States of America and China (Foray, 2014). In doing so, the priority would revolve around funding entrepreneurs and organisations in core regions, and therewith, the preference would be to link entrepreneurs and organisations from core regions together. As a consequence, peripheral regions are often overlooked. Therewith, place-based policy should remain bottom-up and keep in mind the path-dependency of regions which limits the opportunities in peripheral regions. However, in terms of sustainable and inclusive growth, policy-makers and researchers of S3 should seek how to make competitive use of peripheral regions and pick the willing.

6.4. Methodological Suggestions

In this thesis several methodological issues are put forward to develop the entry model and place-based policy further. The suggestions to improve methodology are as following. A general issue is that regional analysis involves aggregations of data in which causality gets lost. For example, measuring the mean of GDP per capita neglects the underlying distribution of economic development within the region. Thus, due to aggregation of data, regions may classify as core whilst parts of the region may be peripheral. These issues lead findings to become rather correlational than causal. Nevertheless, these aggregations of data are necessary to have an eligible amount of patents per region for the analyses in this thesis. Further research could experiment how regression models withstand under smaller aggregations of data (e.g., Functional Urban Areas) through measuring (technological) diversification on basis of surveys. Besides, qualitative research on technological diversification have been widely neglected, thus how the entry model reflects regions in on basis of qualitative judgement is appealing to develop the methodology further.

The entry model is a state-of-the-art model that exist to study technological change with few bias in peripheral regions. Nevertheless, further improvements in the model may lie in introducing the right variables and interaction effects into regression models to deal with heteroskedasticity, applying further model specification to deal with spatial and temporal autocorrelation, and extensive data preparation to deal with sample biases. An improvement to the statistical modelling would be to consider the definition of complementary interregional linkages by Balland and Boschma (2021) to involve the cognitive proximity of interregional collaboration. In other words, interregional collaboration is more effective if the interregional linkages are related to regions with a similar knowledge base. Furthermore, the measure of complementary interregional linkages is specifically at the region-technology level, and thus that definition is less of a one-size fits all approach.

Further research could also take the interregional linkages as the dependent variable of instead of the entry of a new specialisation. Such research could embrace network analysis to apply regression models at the level of interregional linkages (edges) and different types of proximity between regions as the independent variables as explained before. Especially relating different types proximity to interregional linkages would lead a better understanding of why peripheral regions were able to diversify. Therewith, such analysis could investigate the betweenness centrality further.

References

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish journal of emergency medicine*, *18*(3), 91–93.

Asheim, B. T., & Isaksen, A. (1997). Location, agglomeration and innovation: Towards regional innovation systems in norway? *European planning studies*, *5*(3), 299–330.

Balland, P.-A. (2016). Econgeo: Computing key indicators of the spatial distribution of economic activities. *R package*.

Balland, P.-A. (2017). Economic geography in r: Introduction to the econgeo package. *Available at SSRN 2962146*.

Balland, P.-A., & Boschma, R. (2021). Complementary interregional linkages and smart specialisation: An empirical study on european regions. *Regional Studies*, *55*(6), 1059–1070.

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., Jara-Figueroa, C., Petralia, S. G., Steijn, M., Rigby, D. L., & Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities. *Nature human behaviour*, *4*(3), 248–254.

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

Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation. *Progress in human geography*, *28*(1), 31–56.

Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional studies*, *39*(1), 61–74.

Boschma, R., Balland, P.-A., & Kogler, D. F. (2015). Relatedness and technological change in cities: The rise and fall of technological knowledge in us metropolitan areas from 1981 to 2010. *Industrial and corporate change*, *24*(1), 223–250.

Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in italy. *Economic geography*, *85*(3), 289–311.

Boschma, R., Minondo, A., & Navarro, M. (2013). The emergence of new industries at the regional level in s pain: A proximity approach based on product relatedness. *Economic geography*, *89*(1), 29–51.

Breusch, T. S., & Pagan, A. R. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The review of economic studies*, *47*(1), 239–253.

Bryman, A. (2016). *Social research methods*. Oxford university press.

Buenstorf, G., & Klepper, S. (2009). Heritage and agglomeration: The akron tyre cluster revisited. *The Economic Journal*, *119*(537), 705–733.

Caragliu, A., de Dominicis, L., & de Groot, H. L. (2016). Both marshall and jacobs were right! *Economic geography*, *92*(1), 87–111.

Castaldi, C., Frenken, K., & Los, B. (2015). Related variety, unrelated variety and technological breakthroughs: An analysis of us state-level patenting. *Regional studies*, *49*(5), 767–781.

Cortinovis, N., Xiao, J., Boschma, R., & van Oort, F. G. (2017). Quality of government and social capital as drivers of regional diversification in europe. *Journal of Economic Geography*, *17*(6), 1179–1208.

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.

De Noni, I., Orsi, L., & Belussi, F. (2018). The role of collaborative networks in supporting the innovation performances of lagging-behind european regions. *Research Policy*, *47*(1), 1–13.

De Rassenfosse, G., Dernis, H., & Boedt, G. (2014). An introduction to the patstat database with example queries. *Australian economic review*, *47*(3), 395–408.

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.

Eck, N. J. v., & Waltman, L. (2009). How to normalize cooccurrence data? an analysis of some well-known similarity measures. *Journal of the American society for information science and technology*, *60*(8), 1635–1651.

Eder, J. (2019). Innovation in the periphery: A critical survey and research agenda. *International Regional Science Review*, *42*(2), 119–146.

Eder, J., & Trippl, M. (2019). Innovation in the periphery: Compensation and exploitation strategies. *Growth and Change*, *50*(4), 1511–1531.

Eurostat. (2023). Nuts - nomenclature of territorial units for statistics: Background. https://ec.europa.eu/eurostat/web/nuts/background

Fitjar, R. D., & Rodriguez-Pose, A. (2011). Innovating in the periphery: Firms, values and innovation in southwest norway. *European Planning Studies*, *19*(4), 555–574.

Fitjar, R. D., & Rodriguez-Pose, A. (2014). When local interaction does not suffice: Sources of firm innovation in urban norway. In *Regional development and proximity relations* (pp. 195–222). Edward Elgar Publishing.

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

Foray, D. (2014). *Smart specialisation: Opportunities and challenges for regional innovation policy*. Routledge.

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional studies*, *41*(5), 685–697.

Getis, A. (2007). Reflections on spatial autocorrelation. *Regional Science and Urban Economics*, *37*(4), 491–496.

Grillitsch, M., & Nilsson, M. (2015). Innovation in peripheral regions: Do collaborations compensate for a lack of local knowledge spillovers? *The Annals of Regional Science*, *54*(1), 299–321.

Hassink, R., & Gong, H. (2019). Six critical questions about smart specialization. *European Planning Studies*, *27*(10), 2049–2065.

Hidalgo, C. A., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D. F., Morrison, A., et al. (2018). The principle of relatedness. *International conference on complex systems*, 451–457.

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

Iacobucci, D., & Guzzini, E. (2016). Relatedness and connectivity in technological domains: Missing links in s3 design and implementation. *European Planning Studies*, *24*(8), 1511–1526.

Iammarino, S., Rodriguez-Pose, A., & Storper, M. (2019). Regional inequality in europe: Evidence, theory and policy implications. *Journal of economic geography*, *19*(2), 273–298.

Isaksen, A. (2014). Industrial development in thin regions: Trapped in path extension? *Journal of economic geography*, *15*(3), 585–600.

Jalali, A., Tamimi, R. M., McPherson, S. M., & Murphy, S. M. (2022). Econometric issues in prospective economic evaluations alongside clinical trials: Combining the nonparametric bootstrap with methods that address missing data. *Epidemiologic Reviews*, *44*(1), 67–77.

Johnston, R., Jones, K., & Manley, D. (2018). Confounding and collinearity in regression analysis: A cautionary tale and an alternative procedure, illustrated by studies of british voting behaviour. *Quality & quantity*, *52*, 1957–1976.

Klepper, S. (2007). Disagreements, spinoffs, and the evolution of detroit as the capital of the us automobile industry. *Management science*, *53*(4), 616–631.

Kogler, D. F., Essletzbichler, J., & Rigby, D. L. (2017). The evolution of specialization in the eu15 knowledge space. *Journal of Economic Geography*, *17*(2), 345–373.

Krugman, P. R. (1979). Increasing returns, monopolistic competition, and international trade. *Journal of international Economics*, *9*(4), 469–479.

Lau, A. K., & Lo, W. (2015). Regional innovation system, absorptive capacity and innovation performance: An empirical study. *Technological Forecasting and Social Change*, *92*, 99–114.

Le Gallo, J., & Plunket, A. (2020). Regional gatekeepers, inventor networks and inventive performance: Spatial and organizational channels. *Research Policy*, *49*(5), 103981.

Li, D., Heimeriks, G., & Alkemade, F. (2020). The emergence of renewable energy technologies at country level: Relatedness, international knowledge spillovers and domestic energy markets. *Industry and Innovation*, *27*(9), 991–1013.

Mazzucato, M. (2018). Mission-oriented innovation policies: Challenges and opportunities. *Industrial and corporate change*, *27*(5), 803–815.

McCann, P., & Ortega-Argilés, R. (2015). Smart specialization, regional growth and applications to european union cohesion policy. *Regional studies*, *49*(8), 1291–1302.

Newman, M. E. (2002). Assortative mixing in networks. *Physical review letters*, *89*(20), 208701.

Nilsen, T., Grillitsch, M., & Hauge, A. (2023). Varieties of periphery and local agency in regional development. *Regional Studies*, *57*(4), 749–762.

Onsager, K., Isaksen, A., Fraas, M., & Johnstad, T. (2007). Technology cities in norway: Innovating in glocal networks. *European Planning Studies*, *15*(4), 549–566.

Petralia, S., Balland, P.-A., & Morrison, A. (2017). Climbing the ladder of technological development. *Research Policy*, *46*(5), 956–969.

Pugh, R., & Dubois, A. (2021). Peripheries within economic geography: Four “problems” and the road ahead of us. *Journal of Rural Studies*, *87*, 267–275.

Rigby, D. L. (2015). Technological relatedness and knowledge space: Entry and exit of us cities from patent classes. *Regional Studies*, *49*(11), 1922–1937.

Rodriguez-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.

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

Rodriguez-Pose, A., & Fitjar, R. D. (2013). Buzz, archipelago economies and the future of intermediate and peripheral areas in a spiky world. *European Planning Studies*, *21*(3), 355–372.

Steijn, M. (2021). Improvement on the association strength: Implementing a probabilistic measure based on combinations without repetition. *Quantitative Science Studies*, *2*(2), 778–794.

Watson, G. S., & Durbin, J. (1951). Exact tests of serial correlation using noncircular statistics. *The Annals of Mathematical Statistics*, 446–451.

Whittle, A., & Kogler, D. F. (2020). Related to what? reviewing the literature on technological relatedness: Where we are now and where can we go? *Papers in Regional Science*, *99*(1), 97–113.

Whittle, A., Lengyel, B., Kogler, D. F., et al. (2020). Understanding regional branching knowledge diversification via inventor collaboration networks. *Papers in Evolutionary Economic Geography*, *20*.

A
Variable Operationalisation

Table A.1: The operationalisation of variables.

Variable	Model	Operationalisation	Category	Reference
Entry	Dependent	Entry of a technological field i in region r at time t	Binary	*OECD*
Linkages Core	Independent	Number of external co-inventors in core regions s at time t	Discrete	*OECD*
Linkages Periphery	Independent	Number of external co-inventors in peripheral regions s at time t	Discrete	*OECD*
Relatedness Density	Control	Relatedness of a technological field i in region r at time t	Continuous	*OECD*
GDP per Capita	Control	GDP per capita of region r at time t	Continuous	*Eurostat*
Population Size	Control	Population size of region r at time t	Discrete	*Eurostat*
Population Density	Control	Population density of region r at time t	Continuous	*Eurostat*
Absorptive Capacity	Control	Gross R&D expenditure of region r at time t	Continuous	*Eurostat*
Period	Control	Period time t	Categorical	Model

B
Assortative Mixing

Assortative mixing by enumerative characteristics is a measure to compare the actual number A_E and the expected number E_E of linkages between regions their objective (Newman, 2002). The assortativity coefficient r is a measure of homophily and heterophily i.e., the tendency to associate with similar or dissimilar entities. The number of actual linkages between regions of an identical objective is:

$$A_E = \sum_{(i,j)\in E} \delta(c_i, c_j) = \frac{1}{2}\sum_{i,j} a_{i,j}\delta(c_i, c_j) \tag{B.1}$$

where E is the set of linkages in the social network and $a_{i,j}$ is the number of actual linkages between region i and j. The factor one-half accounts for the undirected linkages. The Kronecker delta is a formal definition for regions belonging to identical objectives:

$$\delta(c_i, c_j) = \begin{cases} 0, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases} \tag{B.2}$$

The expected number of linkages between regions of an identical objective is a mathematical estimation as if the objectives are spread randomly over the social network:

$$E_E = \frac{1}{2}\sum_{i,j} \frac{d_i d_j}{2m}\delta(c_i, c_j) \tag{B.3}$$

m is the number of linkages in the social network. Regions i and j yield a degree d_i and d_j respectively. Thus, $\frac{d_i d_j}{m}$ refers to the expected number of linkages between region i and j. The modularity is a measure of difference between the actual and the expected number of linkages:[1]

$$Q = \frac{1}{2m}\sum_{i,j}\left(a_{i,j} - \frac{d_i d_j}{2m}\right)\delta(c_i, c_j) \tag{B.4}$$

whereas the maximum possible modularity is the difference between the total and the expected number of linkages:

$$Q_{max} = \frac{1}{2m}\left(2m - \sum_{i,j}\frac{d_i d_j}{2m}\delta(c_i, c_j)\right). \tag{B.5}$$

All in all, normalising modularity results in the assortativity coefficient r in Equation 3.5.

[1] http://users.dimi.uniud.it/~massimo.franceschet/teaching/datascience/network/assortative.html

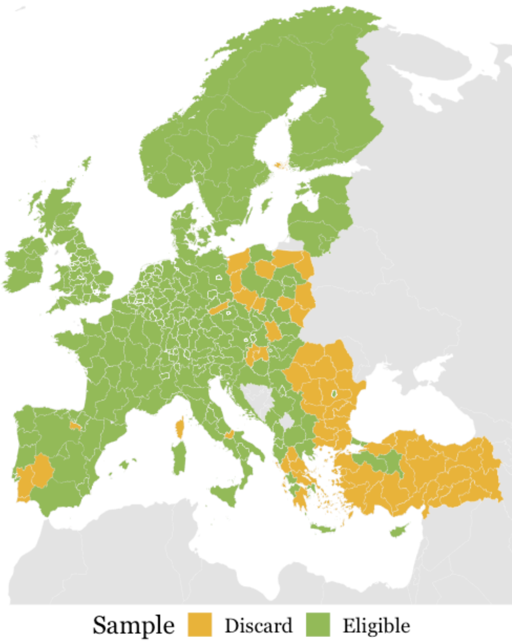
C
Sample Reselection

It is important is to ensure that the quantity of patents in regions is eligible in terms of data validity and outliers, and therewith, to perform computations on. If the number of patents of a region is too low, then the region needs to be discarded. As in Figure C.1a, similar to Balland et al. (2019), if a region exhibits less than 10 fractional patent counts in any of the periods, then the amount of patents in a region could be considered as insufficient. Here, a sufficient amount of patents refers to an eligible region. However, this thesis discards regions on basis of regional eligibility at the country level to draw upon a coherent sample of regions in terms of interregional linkages. In particular, the sample should not consist of regions that would exist in isolation. The selection of regions is restricted to countries in Western Europe (from an iterative perspective) which are available in the dataset of Eurostat (2023). The believe is that this sample of regions in Western Europe links coherently. How the sample of Western regions link is investigated in further analysis.

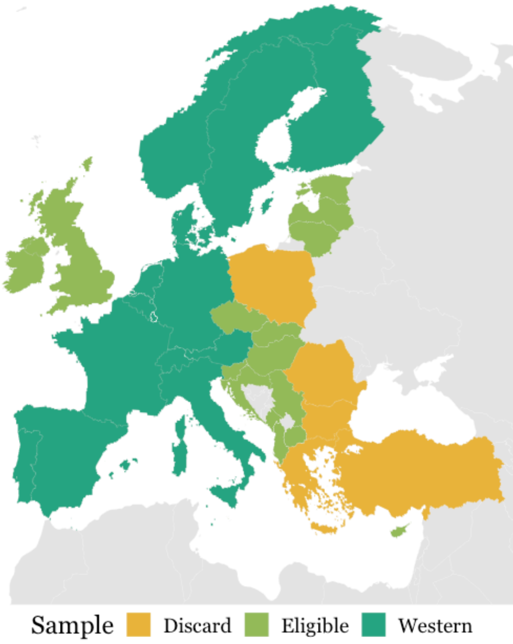
Second is to ensure that the quantity of patents per technological field is sufficient and independent from small absolute numbers. A threshold is set for technological fields to have at least a fractional patent count of 50 over the years $2000 - 2018$. The threshold of excluding technological fields is applied after reconsidering the sample of regions.

The process of sample selection by region is substantiated as follows. The goal is to select countries that fit within a coherent sample of Western Europe. Figure C.1 depicts the overview of the sample. Portugal (PT) is considered as the threshold because this country has the least amount of eligible regions at the country-level within Western European Countries. The countries are discarded if their fraction of eligible regions at the country-level is lower than Portugal (PT). The countries are eligible if their fraction of eligible regions at the country-level is higher than Portugal (PT). The Western European countries are hand-picked in line with Google (2023), the perception to be a coherent sample, and data availability.

The countries Greece (EL), Poland (PL), Slovakia (SK), Bulgaria (BG), Romania (RO), and Turkey (TR) yield the lowest amount of eligible regions over time, and thus these countries are discarded directly. Further, the countries Albania (AL), Cyprus (CY), Estonia (EE), Croatia (HR), Lithuania (LT), Latvia (LV), Montenegro (ME), North Monaco (MK), Malta (MT), Servia (RS), Slovenia (SI), Czechia (CZ), and Hungary (HU) are eligible, yet are discarded, to avoid countries (among their regions) to exist in isolation as of geographic proximity, but also due to cultural proximity. Therewith, the expectation is that interregional linkages among non-Western countries is scarce and would lead to a bias of incoherent collaboration.



(a) The sample by region.



(b) The sample by country.

Figure C.1: The selection of the sample.

The remainder sample of Western European countries includes Austria (AT), Belgium (BE), Swiss (CH), Germany (DE), Denmark (DK), Luxembourg (LU), Netherlands (NL), Norway (NO), Sweden (SE), Italy (IT), Spain (ES), Finland (FI), France (FR), and Portugal (PT). The believe is that these countries are coherent in interregional linkages due to geographic and cultural proximities between those countries. Notable is that Portugal (PT) yields few eligible regions, nevertheless, the country is considered as belonging to Western European countries. Further, due to Brexit no data is available for the United Kingdom (UK) because there is no agreement made on statistical cooperation at the current date (Eurostat, 2023). Lastly, due to regional formations, the regions in Ireland (IE) and Nord-Norge and Trøndelag in Norway (NO) are not compatible in NUTS2 regions between *OECD 2013* and *Eurostat 2023*. The Western European countries, considering all available data, leads to a sample of 176 regions.

The percentage of eligible regions in the sample increases from 77.71% to 91.53% when delimiting to sample of Western European countries. Furthermore, the sample of Western European countries is expected to be coherent in terms of interregional linkages in comparison to the previous sample. The next step is to impute missing years of economic data from *Eurostat 2023*. The last-observation carried forward (then backward) over time for each region is a applied to impute missing data (Jalali et al., 2022).

chapter 4 depicts the core-periphery dichotomy for the new sample. The classification of core and peripheral regions is adjusted relative to the sample of Western countries (by GDP per capita for each year separately). The adjusted classification is not according to the European Cohesion Policy, but still is relevant to the theory and hypotheses.

D
Difference in Entries over the Core-Periphery Dichotomy

The total entries in technological fields variate strongly over the core-periphery dichotomy. The ten most dominant contrasts in technological diversification between core and peripheral regions in $2005-2018$ are listed in Table D.1. The ΔEntry resembles the relative difference of all entries in peripheral regions in contrast to core regions. Note, the ΔEntry deviates further from zero for core regions, meaning that core regions lead by a larger percentage in technological diversification for those technological fields. For example, peripheral (in contrast to core) regions diversify dominantly in preservation of bodies of humans, animals, or plants (A01N), composition of macromolecular compounds (C08L), and fixed or movable closures for openings in buildings, vehicles, or fences (E06B).

Table D.1: Top 10 differences in the entry of technological fields over the core-periphery dichotomy.

CPC4	ΔEntry	Description of the Technological Field
Core		
G01H	-12.33%	Measurement of mechanical vibrations, or ultra- or infrasonic waves
B03B	-12.20%	Separating solid materials using liquids, pneumatic tables, or jigs
F23K	-12.07%	Feeding fuel to combustion apparatus
B25H	-11.96%	Workshop equipment e.g., for marking-out work
B24D	-11.89%	Tools for grinding, buffing, or sharpening
B24C	-11.74%	Abrasive or related blasting with particulate material
B64F	-11.71%	Ground or aircraft-carrier-deck installations
E05G	-11.35%	Safes or strong-rooms for valuables; Bank protection devices
A63F	-11.19%	Card, board, or roulette games; Indoor games; Video Games
B42F	-11.14%	Sheets temporarily attached together; Filing appliances
Periphery		
H04L	5.19%	Transmission of digital information
A61F	5.20%	Filters implantable into blood vessels; Prostheses; Stents
C08G	5.76%	Macromolecular compounds obtained otherwise than by reactions
A61Q	5.96%	Specific use of cosmetic or similar toiletry preparations
C07C	5.99%	Acyclic or cabrocyclic compounds
C07K	6.09%	Peptides
A61K	6.41%	Preparations for medical, dental, or toiletry purposes
E06B	6.76%	Fixed or movable closures for openings in buildings, vehicles, or fences
C08L	7.16%	Composition of macromolecular compounds
A01N	7.87%	Preservation of bodies of humans, animals, or plants