DISSERTATION

zur Erlangung des akademischen Grades der

Doktor-Ingenieurin (Dr.-Ing.) der Fakultät Raumplanung der

Technischen Universität Dortmund

Vorgelegt von

Vanessa Hellwig, M.Sc.

aus: Moers, Deutschland

Tag der mündlichen Prüfung: 11.10.2022

Topic

Transformation, Digitisation and the Geography of Knowledge

—

Essays on the Spatial Distribution of the Digital Economy in Germany

Supervisors:

Prof. Dr. Christiane Hellmanzik, TU Dortmund University, Germany
Prof. Dr. Stefan Siedentop, TU Dortmund University, Germany
Prof. Dr. Tessa Flatten, TU Dortmund University, Germany

Acknowledgements

My special thanks go to Prof. Dr. Christiane Hellmanzik. I am grateful for her guidance, feedback and for sharing her enormous knowledge with me. Even more importantly, I am indebted for her trust in my abilities and for her always seeing the person before the student.

I would also like to take the opportunity to thank Prof. Dr. Stefan Siedentop for immediately agreeing to be my second supervisor. His extensive expertise, valuable comments and suggestions contributed a lot to this dissertation. Further, I thank Prof. Dr. Tessa Flatten for her support and her flexibility, and for always finding time in a tight schedule.

I had the pleasure to work in a great team environment. Dr. Sara Mitchell gave great guidance and support during very early research phases. I thank Dr. Lukas Kuld for being a rough, but likable criticizer and for making pre-covid office hours a lot more entertaining. I am very grateful to be working with Felix Dornseifer and Maren Kaliske. Thank you for creating safe spaces to ask questions. Especially Maren's friendship in- and outside the academic sphere were incredibly valuable for me. I thank Jens Kirsten, Jan Lordieck, Hannah Bellgardt and Adrian Hinse for excellent research assistance during the past years.

Moreover, I would like to thank my friends, especially Hanna Seydel, Isabelle Wachter and Lukas Naumann not only for being my intellectual hook to the planning department and PhD sparring partners, but for ten years of friendship. I thank my parents Monika and Rainer for their unconditional support and trust. My siblings Denise and Christian have been among the most important people in my life and I am unbelievable proud of them. Finally, I thank my beloved Jens for everything. With you, things are easy.

Vanessa Hellwig, M.Sc. Dortmund, 27 October 2022

Executive Summary

This dissertation provides empirical analysis and findings on patterns of the spatial development of the digital industry in Germany over time. In addition, the results allow conclusions on the spatial dimension and channels of agglomeration economies and the dissemination of knowledge captured by entrepreneurship in a spatial context. The thesis consists of six main chapters: The first and second chapter provide a short introduction to the topic including the relationships between innovation, digitisation, the geography of knowledge and entrepreneurship. This is followed by three empirical contributions in chapters three, four and five at the interface between spatial questions relevant for urban and regional planning by simultaneously applying theory and methodologies derived from urban economics. The analysis allows to gain knowledge on the spatial development and determinants of the digital industry in Germany being of major interest for policy makers to foster local growth. The last chapter provides a summary of the main findings and concluding remarks.

In this dissertation, the quantitative research setup builds on a unique, tailor-made geo-coded firm-level panel dataset of digital firms. The uniqueness lies in the precise tracking of firm locations over time. For each contribution, the data is aggregated and applied in three distinct spatial scales that are NUTS 3 regions (counties), LAU Regions (municipalities) within urban labour market regions and 1x1 km^2 grids within Hamburg, Berlin and Munich.

To begin with, the first empirical contribution analyses and compares firm birth and relocation patterns of digital firms between 2008 to 2017 on county level (NUTS 3) to understand differences in location choices within firms' life cycles. I match data on 107,321 firms aggregated on county level with administrative data to capture local characteristics. My findings are largely in line with the literature on the spatial patterns for knowledge-intensive industries: The regional knowledge base stemming from universities and co-located similar firms is a key-determinant for digital firm birth. In order to analyse the relocation of young firms (<10 years), the contribution is the first to employ a fixed-effects gravity model, the workhorse model for international trade

and FDI. The model uses aggregated relocation flows instead of individual firm decisions as in discrete choice models, the go-to model in the literature on relocation of firms. The findings indicate, in line with the predictions of the gravity model, that relocation flows are highest between contiguous counties and over short distances as relocation costs increase with distance. This result implies a regional persistence of entrepreneurship. This indicates that large spatial shifts of the maturing industry are not to be expected.

In the second contribution in chapter four, the units of analysis are municipalities (LAU regions) within urban labour market regions. This fourth chapter draws on an interdisciplinary approach by investigating industry dynamics in mono- and polycentric city regions between 1995-2017. Scholars in economics usually assume a standard "core - urban periphery" logic. Possible differences in the core-urban-periphery dynamic in monocentric and polycentric urban area are highly relevant for the planning community. Thereby, scholars in the planning community mostly focus on population distributions instead of firm behavior. This is why I estimate what characteristics of municipalities are conducive to attracting firm birth while also including characteristics of the next core city. The analysis is done for mono-and polycentric areas individually to detect possible differences. The analysis indicates two main results: First, small municipalities close to core cities gain advantages over their equally small neighbours by hosting a university and from population growth. Second, the regional pattern of the digital industry is shaped by the morphology and digital sector in the closest core cities: Municipalities in monocentric urban regions profit from sharing (population growth) and general knowledge from universities, while municipalities in polycentric urban regions are affected by industry-specific externalities that is an above-average growth in the share of firm birth in their closest urban cores.

The third contribution takes on a micro-geographical approach by linking location choices of digital firms within cities to knowledge institutions (Higher Education Institutions and research institutes) on a 1x1 km^2 grid level. Empirical papers on such small spatial scales are still sparse, mostly due to limited availability of high quality data. The contribution lies, apart from the spatial unit of analysis, in the combination of the literature of agglomeration externalities and the differentiated knowledge base approach. The latter has predominantly been investigated using qualitative methods. Results reveal strong clustering of firms within cities close to universities. However, within-industry knowledge spillovers decay more rapidly over distance than industry-institution spillovers. Next to spatial proximity, digital firms favor economic proximity. That is knowledge contents which can directly be transferred into digital products. The knowledge sources are mainly similar firms and design-schools containing highly context specific tacit knowledge that is hard to retrieve from further away.

Overall, the empirical contributions indicate several main takeaways. Digital firm birth is an urban phenomenon. Young firms are dependent on outside resources and favor thick labour markets with highly educated human capital and knowledge sources. These knowledge sources are Higher Education Institutes and similar firms. Hence, policies targeting homogeneous digital clusters based on co-location are conducive to the local growth of the digital sector. This is because localized spillovers (similar knowledge) decay more rapidly with distance than urbanized (general) spillovers. Further, there is no spatial displacement of the industry over long distances. This means that a locations' own regional entrepreneurial capital is of fundamental importance for the digital sector.

Kurzzusammenfassung

Die vorliegende Dissertation liefert Erkenntnisse zur räumlichen Entwicklung der digitalen Wirtschaft in Deutschland. Darüber hinaus lassen die Ergebnisse Rückschlüsse auf die räumliche Dimension von Agglomerationsvorteilen und Kanäle von Wissensdiffusion zu, welche durch Unternehmertum in einem räumlichen Kontext erfasst werden. Die Dissertation besteht aus sechs Kapiteln. Das erste und zweite Kapitel enthalten eine kurze Einführung in das Thema, einschließlich der Beziehungen zwischen Innovation, Digitalisierung und der Geografie von Wissen und Unternehmertum. Es folgen drei empirische Beiträge in den Kapiteln drei, vier und fünf. Die empirischen Beiträge bewegen sich an der Schnittstelle zwischen stadt- und regionalplanungs relevanten räumlichen Fragen, unter Berücksichtigung von Theorien und Methoden aus der Volkswirtschaftslehre. Das letze Kapitel legt eine Zusammenfassung der Ergebnisse und Schlussfolgerungen dar.

Die empirischen Beiträge basieren auf einem einzigartigen, maßgeschneiderten, geo-kodierten Panel-Datensatz zu digitalen Unternehmen. Die Einzigartigkeit liegt in der präzisen Verfolgung der Firmenstandorte über die Zeit. Die Analyse der Firmenstandorte ermöglicht es, Erkenntnisse über die räumliche Entwicklung und die Determinanten der digitalen Industrie in Deutschland zu gewinnen. Dies ist für politische EntscheidungsträgerInnen von großem Interesse, um lokales Wachstum zu fördern. Darüber hinaus bietet die Arbeit Einblicke in die räumliche Dimension wirtschaftlicher Mechanismen, d.h. Agglomerationsökonomien und die räumliche Diffusion von Wissen und Unternehmertum.

Für die Beiträge werden die Daten auf drei verschiedene räumliche Ebenen aggregiert: NUTS 3 (Kreise und kreisfreie Städte), LAU Regionen (Gemeindeverbände) innerhalb städtischer Arbeitsmarktregionen und die innerstädtische Grid-Ebene (1x1 km^2) innerhalb von Hamburg, Berlin und München.

In dem ersten Beitrag werden räumliche Muster von Gründungsaktiviäten und Firmenumzügen zwischen 2008 und 2017 auf Kreisebene miteinander verglichen. Dies dient dem Verständnis

von Unterschieden in der Standortwahl innerhalb der Lebenszyklen von Unternehmen. Für Firmengründungen werden Daten zu 107.321 digitalen Firmen mit administrativen Daten zusammengeführt, um lokale Merkmale zu erfassen. Die Ergebnisse stimmen weitgehend mit der Literatur zu den räumlichen Mustern in wissensintensiven Branchen überein: Die regionale Wissensbasis - die von Universitäten und ähnlichen Unternehmen am gleichen Standort stammt - ist ein wichtiger Schlüsselfaktor für die Gründung digitaler Unternehmen. Für die Analyse von Standortverlagerungen junger Unternehmen (<10 Jahre) wird erstmals ein Gravitationsmodell verwendet, das bewährte Modell für internationalem Handel und ausländische Direktinvestitionen. Das Modell verwendet aggregierte Verlagerungsströme anstelle von individuellen Unternehmensentscheidungen wie in diskreten Wahlmodellen, dem Standardmodell in der Literatur zu Unternehmensumzügen. Die Ergebnisse zeigen in Übereinstimmung mit den Annahmen des Gravitationsmodells, dass die Verlagerungsströme zwischen benachbarten Kreisen und über kurze Distanzen am höchsten sind, da die Verlagerungskosten mit der Entfernung zunehmen. Dies zeigt eine regionale Persistenz des Unternehmertums auf. Die Ergebnisse deuten daher daurauf hin, dass eine goße Verlagerung in der Industrie nicht zu erwarten ist.

Der zweite Beitrag betrachtet die Entwicklung der Digitalbranche in Gemeindeverbände (LAU Regionen) innerhalb städtischer Arbeitsmarktregionen. Dieser Beitrag stützt sich auf einen interdisziplinären Ansatz, indem er die Branchendynamik in mono- und polyzentrischen Stadtregionen zwischen 1995 und 2017 untersucht. Wirtschaftswissenschaftliche Arbeiten gehen in der Regel von der Standardlogik "ein Kern - städtische Peripherie" aus, während mögliche Unterschiede zwischen mono- und polyzentischen Stadtgebieten für die Planungsgemeinschaft von großer Bedeutung sind. Die Planungsgemeinschaft konzentriert sich meist auf die Bevölkerungsverteilung und nicht auf das Verhalten der Firmen. Aus diesem Grund zielt die ökonometrische Schätzung darauf ab, welche Merkmale von Gemeinden für die Ansiedlung von Unternehmen förderlich sind, unter Berücksichtigung der Merkmale der nächsten Kernstadt. Die Analyse wird für mono- und polyzentrische Gebiete einzeln durchgeführt, um mögliche Unterschiede zu ermitteln. Die empirischen Ergebnisse des Beitrags zeigen einerseits, dass kleine Gemeinden in der Nähe von Kernstädten durch die Ansiedlung einer Universität und durch Bevölkerungswachstum Vorteile gegenüber ihren vergleichbar kleinen Nachbarn haben. Andererseits wird das regionale Muster der digitalen Industrie wesentlich durch die Morphologie und dem digitalen Sektor in den nächsten Kernstädten geprägt: Gemeinden in monozentrischen Stadtregionen profitieren von Bevölkerungswachstum und Wissen von Universitäten, während Gemeinden in polyzentrischen Stadtregionen von branchenspezifischen externen Effekten betroffen sind, d.h. einem überdurchschnittlichen Wachstum des Anteils der Unternehmensgründungen in den nächstgelegenen Stadtkernen.

Der dritte Beitrag verfolgt einen mikrogeografischen Ansatz, indem er die Standortentscheidungen digitaler Unternehmen innerhalb von Städten mit Wissenseinrichtungen verbindet. Empirische Arbeiten in kleinen räumlichen Maßstäben sind selten, aufgrund der begrenzten Verfügbarkeit hochwertiger Daten. Der Beitrag liegt, abgesehen von der räumlichen Analyseeinheit, in der Kombination der Literatur zu Agglomerationsexternalitäten und dem Ansatz der differenzierten Wissensbasen. Letzterer ist überwiegend mit qualitativen Methoden untersucht worden. Daher werden in der Analyse die Gründungsmuster digitaler Unternehmen in der Nähe von unterschiedlichen Hochschulen (Forschungsuniversitäten, Fachhochschulen und Kunst- Musik und Designhochschulen) untersucht sowie die Wissensinhalte der ansässigen Fakultäten im Zusammenhang mit der 'ökonomischen Nähe' zur Digitalindustrie. Die Ergebnisse der Analyse zeigen eine starke Clusterbildung von Unternehmen in Städten in der Nähe von Universitäten. Der Wissensspillover innerhalb einer Branche nimmt jedoch auf der räumlichen Ebene schneller ab als der Spillover zwischen Branchen und Institutionen. Zudem bevorzugen digitale Unternehmen die wirtschaftliche Nähe zu Wissensquellen, d.h. Wissen, das ohne größere Umwege in digitale Produkte umgesetzt werden kann. Bei diesen Quellen handelt es sich vor allem um ähnliche Unternehmen und Designhochschulen, die über sehr kontextspezifisches implizites Wissen verfügen, auf das aus größerer Entfernung nur schwer zugegriffen werden kann, aber zur Marktfähigkeit der Produkte beiträgt.

Insgesamt lassen sich aus den empirischen Beiträgen mehrere Kernaussagen ableiten. Das Gründungsgeschehen digitaler Unternehmen ist ein städtisches Phänomen. Junge Unternehmen sind auf externe Ressourcen angewiesen und bevorzugen dichte Arbeitsmärkte mit hochqualifiziertem Humankapital und Wissensquellen. Diese Wissensquellen sind Hochschulinstitute und andere Unternehmen deren Geschäftemodell auf einer ähnlichen Wissensbasis fundiert. Daher sind politische Maßnahmen, die auf homogene digitale Cluster auf der Grundlage der Kolokation abzielen, für das lokale Wachstum des digitalen Sektors zuträglich. Dies liegt daran, dass lokalisierte Spillover (ähnliches Wissen) mit der Entfernung schneller abnehmen als urbanisierte (allgemeine) Spillover. Außerdem gibt es keine Verlagerung der Branche über große Entfernungen, d.h. das eigene regionale Unternehmerkapital ist für den digitalen Sektor von grundlegender Bedeutung.

Eidesstattliche Versicherung

Gemäß § 11 der Promotionsordnung der Fakultät Raumplanung der Technischen Universität Dortmund erkläre ich folgende Punkte:

1. Bei der eingereichten Dissertation zu dem Thema „Transformation, Digitisation and the Geography of Knowledge“ handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.

3. Die Arbeit oder Teile davon habe ich bislang nicht an einer Hochschule des In- oder Auslands als Bestandteil einer Prüfungs- oder Qualifikationsleistung vorgelegt.

4. Die Richtigkeit der vorstehenden Erklärungen bestätige ich.

5. Die Bedeutung der eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

Ort und Datum	Vanessa Hellwig, M.Sc.

Contents

List of Figures

List of Tables

List of Abbreviations

BBSR	Bundesinstitut für Bau-, Stadt- und Raumforschung
BMWK	Bundesministerium für Wirtschaft und Klimaschutz
CBD	Cental Business District
EU	European Union
FDI	Foreign Direct Investment
GDP	Gross Domestic Product
HEI	Higher Education Institution
ICT	Information and Communication Technologies
IT	Information Technology
MUR	Monocentric Urban Region
NB	Negative Binomial
OLS	Ordinary Least Squares
OSM	OpenStreetMap
PPML	Pseudo Poisson Maximum Likelihood
PUR	Polycentric Urban Region
R&D	Research and Development
RWI	Leibniz-Institut für Wirtschaftsforschung
SME	small and medium-sized enterprises
STEM	Science, Technology, Engineering, and Mathematics
UAS	University of Applied Science

Chapter 1

Introduction

The main objective of this dissertation is to analyse the spatial development of the digital industry and its determining factors in Germany. By investigating location patterns of firm birth and relocation, the results provide insights into agglomeration effects and spatially bounded knowledge flows, that is the geography of knowledge, beyond the mere description of the industry development. The following chapter gives a general introduction to the relationships of knowledge creation, knowledge exchanges' links to innovation, the digital sector in Germany and why spatial patterns of the latter are of great importance for urban planners as well as economists.

1.1 Transformation, Digitisation and Innovation

Urban structures exhibit a unique blend of persistence and transformation. For some, the foundations were laid centuries ago and became firmly anchored over time. For example, many European cities are still shaped by medieval street layouts. However, cities are equally characterised by economic and societal changes, keeping up with modern design and technology. In recent centuries, there has been a change from the industrial city to a service based community and finally, to a knowledge-based society. Recently, sectors of the knowledge-based economy such as the digital sector became increasingly important for urban development. This goes hand in hand with an ever better educated population seeking to live in cities due to their professional opportunities as well as the host of amenities. During this development, technology and digital tools evolved into a major role in a knowledge-based economy. As a result, well proven political strategies to support local economies e.g. in recruiting existing industrial players as regional growth engines, have been replaced by strategies to foster entrepreneurs that grow into successful drivers of the economy (Henderson and Weiler, 2010). Thus, understanding the companies, technologies and innovators which drive urbanisation as well as digitisation and their allocation strategies is key to yield policies for sustainable development while providing an environment which is conducive to economic activity and future competitiveness.

In this context, innovations are the origin of transformation. Innovations can be stimulated and evoked by different conditions. They are used to solve acute problems by opening new possibilities, such as technological progress, or by recombining existing ones (Taalbi, 2017). Innovations are essential to maintain long-term competitiveness and growth. However, this is true not only for businesses, but also for cities and their associated urban economies. Innovations are also well described by the dynamics of endogenous growth models (Romer, 1990).

On a more abstract level, regional policy markers are given the opportunity to support the local economy in an innovative manner by supporting factors and sectors which influence urban and regional innovation systems. These factors are, among others, firm-specific capabilities and leader firms, research and education infrastructure, entrepreneurship and local markets for high-tech products (Van Winden et al., 2014). It becomes clear that digitisation and innovation processes take part in economic, social and urban transformation. They play an important role considering the future competitiveness of firms, social interactions in everyday life and tomorrows cities.

For that reasons, this dissertation investigates the digital sector as a key driver for such developments described above. The aim of this dissertation is to provide empirical findings on the evolution of the digital industry and its determining factors in Germany. The following section provides a short conceptual overview on the theory of technological change, innovation and entrepreneurship.

1.1.1 Linking Technological Change, Innovation and Entrepreneurship

Innovations are the origin of technological progress and drivers of societal mega-trends such as digitisation. This is why a deep understanding of the origin of innovation and novel knowledge is key to understanding the evolution of the digital industry. Innovations are subject of research in many disciplines, ranging from economics and other social sciences to natural sciences and engineering. Therefore, there is much variation in definitions. Economic innovation research is significantly influenced by the Austrian economist Schumpeter (1934). His perception of innovations are from a behavioral theoretical perspective and places pioneering entrepreneurs in the focus of consideration (Schumpeter, 1934; Fritsch, 2017).

Schumpeter (1934) defines innovation as the creation of new knowledge or the implementation of existing knowledge in a new way and conceptually distinguishes innovations from inventions. An invention, the pure idea of something new and driver of technological progress, becomes an innovation through market introduction or commercialisation (Schumpeter, 1934; Carlino and Kerr, 2015). Thus, the focus of this definition lies on the economic significance of the innovation. The Schumpeterian view sees entrepreneurship as the bottleneck of economic growth by commercialising new knowledge. Indeed, the most significant innovations have predominantly been introduced by small entrepreneurial enterprises (Baumol, 2004; Fritsch, 2017). This is why measuring firm birth activity is a commonly used proxy for innovation and entrepreneurship in empirical studies.

Technological Progress and Economic Growth Theory

The outstanding role of technological progress and innovation has also been part of economic growth theory, that is modelling continuous economic growth and providing theoretical explanations why some countries are rich and others remain poor (Jones, 2002).

The first link between technological progress and economic growth was made by Solow (1956), setting a Nobel Prize winning cornerstone for neoclassical growth models. In essence, his model assumes that accumulation of capital, a growing labour force and externally driven technological change lead to economic growth (Solow, 1956; Jones, 2002). However, technological progress is exogenous, i.e. an external effect originating from automatic and unmodeled improvements in technology. Accordingly, the technology available for firms is unaffected by the actions of the firms, including research and development (Jones, 2002).

Later on, Romer (1990) and Lucas Jr (1988) succeeded in including this exogenous technological progress into the endogenous growth theory, by particularly acknowledging the creation of new knowledge in explaining ongoing economic growth. Here, investments in knowledge and human capital generate positive externalities that increase not only the income of the investor but also that of other actors. Externalities can occur within a sector or between different sectors, leading to macroeconomic synergy effects. These effects enable sustained growth.

Lucas Jr (1988) highlights the presence of human capital by linking 'learning by doing' and competitive advantage. Highly educated and specialized workers are better able to absorb new knowledge from others. The acquisition of knowledge relates to urban contexts, as the high density of people and companies in cities create a fertile environment in which ideas are transferred more quickly. Cities thus support the transfer of knowledge in the form of spillover effects (Carlino and Kerr, 2015).

Romer (1990) focuses on intentional investments in Research and Development (R&D). Technological progress and change is the central driver, where a set of technological opportunity is created by investments in new knowledge by profit-maximizing firms. However, knowledge acquired in R&D cannot be fully protected by the researching firms (only partly and temporarily via patenting) and can diffuse in space, that is spillover. Due to non-rivalry of knowledge, other companies benefit from the knowledge gained by one company, which makes knowledge a positive externality (Feldman and Storper, 2018).

Knowledge Spillover Theory of Entrepreneurship

More recently, Acs et al. (2009) introduced the "Knowledge Spillover Theory of Entrepreneurship". It combines the micro-economic foundations of endogenous growth and entrepreneurship by shifting the unit of analysis from firms to individuals with new knowledge endowments. An individual will start a new business if the expected value of a piece of knowledge is higher for the individual than for a decision maker within an incumbent firm or university (Acs et al., 2009; Audretsch et al., 2008). By commercialising an idea independent from the incumbent organisation via the creation of a new firm, the entrepreneur channels the knowledge spillover into enhanced economic performance. Accordingly, a new idea can evolve in an incumbent firm or organisation as well as in startups. Further, Baumol (2004) states that technological breakthroughs come from predominantly small firms, particularly in the software-sector. In essence, the knowledge spillover theory of entrepreneurship focuses on entrepreneurial behavior within the context of knowledge spillovers (Acs et al., 2013).

The theory poses a foundation for scientist to measure innovations via entrepreneurship and firm birth, because knowledge flows are hard to measure by their very nature. Knowledge originating in formal R&D processes and associated flows can be measured by patent citations (Feldman and Storper, 2018; Henderson and Weiler, 2010). However, the measurement of patents for innovation has several limitations. First, patents are predominately used for goods, as services and intangible goods are hard to patent. In addition, not all patentable goods are actually patented for two main reasons: Large companies do not want to disclose their new research results (Rammer, 2002), and small companies often refrain from filing patent applications for reasons of cost and high legal expenses (Tura et al., 2008). Another argument against using patent citations is that a patent application does not measure the economic value of the innovation. It is therefore possible that a patent application is filed, but the invention does not become an innovation, i.e. the product is never launched on the market (Tura et al., 2008). Due to this shortcomings, measuring the number of new firms entering the market has become a widely used approach in the literature on entrepreneurship (Glaeser et al., 2010; Fritsch and Storey, 2014; Buczkowska and de Lapparent, 2014; Glaeser et al., 2015).

As for the reasons elaborated above the theoretical backbone of this dissertation is the endogenous growth theory and the knowledge spillover theory of entrepreneurship. The exploitation of new knowledge via entrepreneurship can be measured by firm birth activity following Schumpeter's definition of an innovation. This is why I use registration of new, knowledge-intensive firms as a central measure in this dissertation. The digital industry is particularly well suited for this approach, as firms need little physical inputs and sunk costs are low for setting up a business. Further, for many digital business models, outlets are not limited to a local market, but easy to distribute over relatively long distances, mostly in a national market. Hence, for ensuring competitiveness, the digital industry is disproportionately characterised by innovations and the commercialisation of new knowledge (Baumol, 2004; Henderson and Weiler, 2010).

1.1.2 Agglomeration and the Geography of Knowledge

The inclusion of the spatial dimension is a logical step when assuming that not only firm-internal processes and procedures contribute to innovations, but capabilities, external conditions and spillovers are related to innovation processes (Chesbrough, 2003). The basic question of the theoretical debate at the interface of the economic and geographical literature is how entrepreneurship, that is new innovative firms and company-internal scale-effects, e.g. growth, are generated by company-external agglomeration effects or externalities (Van Oort and Atzema, 2004). The following section shortly introduces concepts of agglomeration and the geography of entrepreneurship, innovation and knowledge.

Like no other region, California's Silicon Valley south of San Francisco is known for innovation and pioneering entrepreneurship. Only a stone's throw away from each other, the most famous of all technology companies are located: Apple and Google, among the most valuable companies in the world, Meta, Intel and Cisco, to name just a handful of the big players. Additionally, there are numerous startups that are run out of a garage today but could be worth billions of dollars tomorrow (Weidenbach, 2022). The Californian 'Valley of the Future' is undoubtedly the best-known innovative spot on this earth and provides a vivid illustration of the spatial concentration of innovation and entrepreneurship. Instead of a death of distance, that is overcoming the geographical barrier to the diffusion of information, as predicted by researchers in the early stages of the internet, reality turned out to be the opposite, at least for the digital industry itself (Venables, 2001).

It is well established in the literature, that population and economic activity are spatially concentrated (Carlino and Kerr, 2015). One of the main reasons why firms tend to locate and cluster in cities is because they derive advantages from the agglomeration of population and firms. The main benefits are the pooled labour market (matching), access to specialized suppliers (sharing) and benefits from knowledge spillovers (learning) (Krugman, 1991; Duranton and Puga, 2004; Armington and Acs, 2002). These agglomeration externalities contribute to an increased productivity of the firms, resulting in higher wages for employees (Combes et al., 2012). Especially for knowledge-intensive firms, these advantages outweigh the costs (e.g. higher land prices) and disadvantages from negative agglomeration externalities (e.g. pollution and traffic congestion). This is in contrast to location requirements of industrial production, which demands relatively low cost industry parks and low wages along with a pool of qualified, but not necessarily highly qualified workers (Audretsch et al., 2012).

The learning channel and spillovers are particularly important for digital companies, as new knowledge created in both public and private knowledge institutions and similar firms manifests itself in additional firm birth (Acs et al., 2009; Audretsch et al., 2008) as laid out above. Additionally, young and small firms usually face limited internal competencies, for example constraints in the technical ability of e.g. employees (Schartinger et al., 2001). Due to these constraints, they are likely to run into problems during innovation processes. Because internalizing the needed knowledge is highly costly, they rely on outside resources.

To clarify the understanding of the transmission channels of co-agglomeration forces, the literature distinguishes localisation economies, that is, economies of agglomeration within the same industrial sector, from urbanisation economies, that is, economies of agglomeration between sectors (Glaeser et al., 1992).

Localisation describes co-location of similar firms in close proximity (Porter, 1990) with firms mainly benefiting from employment advantages such as specialized employees and a lower probability of labour shortages (Krugman, 1991). Particularly important benefits for software firms are knowledge spillovers accruing from inter-firm cooperation as well as fluctuation of employees as carriers of knowledge (Trippl et al., 2009). Moreover, co-location of related industries fosters entrepreneurship by lowering costs of starting a business for individuals and enabling better access to a more diverse range of inputs and complementary goods (Glaeser and Kerr, 2009).

Urbanisation effects refer to the benefits of diversity and density of amenities such as public infrastructure which cities typically offer (Jacobs, 1969). Next to the size of the labour market, universities, research institutes and other knowledge- and research-related activities facilitate knowledge spillovers. There is an interdependence between existing knowledge institutions and industry players which is conducive to entrepreneurship within the spatial reach of such knowledge spillovers (Anselin et al., 1997; Bade and Nerlinger, 2000; Trippl et al., 2009; Fritsch and Aamoucke, 2013). Thus, a vital regional knowledge base is more likely to be bigger in cities making cities particularly attractive for young, innovative firms (Acs et al., 2009).

The spatial reach of the knowledge spillovers are thus central in understanding local and regional industry developments. In general, knowledge comes in many different shapes. Codified knowledge is laid out in written form, for example via academic publications, books or patents. Therefore, it is physically bounded, but due to the internet, its usage has become widely independent from the location of the user (Acs and Varga, 2005).

In many cases however, knowledge is not codified, for example novel, not fully developed knowledge, or context specific practical knowledge. Transmitting this type of highly tacit knowledge requires face-to-face contacts and social interaction. This is why knowledge spillovers work on small spatial scales and are locally bound (Van Soest et al., 2006; Larsson, 2014; Jang et al., 2017; Rammer et al., 2020; Roche, 2020).

Thus, regional differences in economic growth fueled by innovation and technological change can be explained by the regional knowledge base consisting of knowledge institutions, that is universities and research institutes as well as incumbent firms. This is also why regional development patterns have been proven to be very persistent over time, and explains ongoing regional disparities (Fossen and Martin, 2018; Stuetzer et al., 2021). Another distinctive factor for the regional persistence is the presence of individuals willing to start a business that is entrepreneurship capital (Audretsch and Keilbach, 2004). Besides the human capital, some regions inherit a deeply rooted social acceptance to encourage and support startup activities through norms and values as well as strong formal and informal networks. Therefore, the local culture serves as a driver for entrepreneurial capital (Audretsch et al., 2008).

This dissertation draws on linking the agglomeration externalities, urbanisation and localisation dimension of entrepreneurship as measured with digital firms in Germany. The contributions presented in the dissertation include empirical analyses on the spatial development patterns of the digital industry and its determining factors. Based on the theory on entrepreneurship, the ideas that lead to start a digital business are likely to originate in incumbent firms and knowledge institutions - that is universities and research institutes. The ex-ante expectation for the location behaviour of the firms is a strong tendency to locate in urban areas to exploit agglomeration externalities from sharing, matching and learning in particular. By that, the thesis aims to provide findings on the geography of knowledge.

1.2 The Digital Economy in Germany

The next section provides an overview on the digital industry in Germany, its contributions to the general economy and the implications this has for cities.

Firms in the digital sector offer technologies and services for data processing and communication. Thereby, Information and Communication Technologies (ICT) services complement to almost all other industries, all needing digitisation in order to keep their competitiveness.

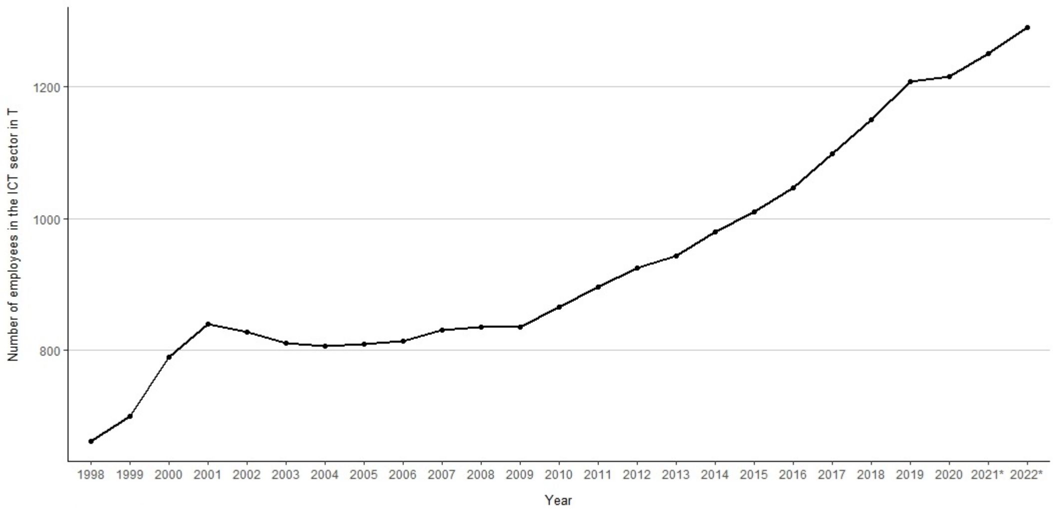
The digital industry is characterised by its strong innovative nature. According to the Federal Ministry of Economic Affairs and Climate Action (Bundesministerium für Wirtschaft und Klimaschutz (BMWK)), 85 % of the firms in the digital sector introduced a new or notably improved product in 2018 (Bertschek et al., 2020). This renders the most innovative of all economic sectors, an economic powerhouse that is politically framed as 'key sector' (Bayer et al., 2022).

Next to, or due to the innovative strength of the sector, it grows faster than average and contributes significantly to the Gross Domestic Product (GDP). The gross value added of the ICT sector rose from 117 to 120 billion euros from 2018 to 2019. Thereby, the sector generates 5.1 % of the value added to the economy as a whole, putting it in eighth place in a sector comparison (Bertschek et al., 2020). Generating digital firms is highly attractive for cities as a key challenge for future competitiveness. Those attracted firms contribute to local growth, demand high skilled labour, do not imply intensive land consumption and do not generate negative environmental effects (Berger and Frey, 2016), as 97 % in ICT startups are ICT services (Bertschek et al., 2020).

The digital sectors' promising contributions to the economy are subject to public funding. Due to funding from many public agencies at several policy levels (e.g. local, regional and national) and a lack of comprehensive data on funding amounts, the precise amount of public funding for ICT firms remains unclear. However, for example the public-private partnership investment fund “High-Tech Gründerfonds” explicitly invests in innovative, high-tech startups, with a budget of about 900 million euros (1.1 billion US $, (High-Tech Gründerfonds, 2021)).

In addition, the digital industry like the German economy in general, is strongly characterised by small companies. In 2019, more than 90 % of firms had less than ten employees and only 0.44 % of the firms employ more than 250 people (Destatis, 2021). The average ICT company employed twelve people, which is slightly above the German average of ten employees per company (Bertschek et al., 2020). This emphasizes the reliance on outside resources for ICT companies in innovation processes as laid out above. Further, hand in hand with growth in firms, there is a growth in employees in the sector (see Figure 1.1).

Figure 1.1: ICT Employment



Notes: Figure 1.1 shows the number of employees (employed and self-employed) in thousands in the ICT and consumer electronics sector in Germany from 1998 to 2022, as of the end of each year. * Forecast/estimate. Datasource: Bitkom (2022).

Figure 1.1 shows a clear upward-trend in labour demand in ICT and the consumer electronics sector. For the ICT sector, this implies the need for high-skilled, highly trained and educated human capital. According to the job center, of 54,000 job vacancies in 2018, 48 % were directed to experts whose qualifications require at least a four-year degree in higher-education. Almost one in three job postings was directed at specialists with a job profile that is typically supported by three years of higher education (such as a bachelor’s degree). One out of every four to five positions should be filled by a specialist with vocational training (Bundesagentur für Arbeit, 2019). This highlights the transition to the knowledge economy in Germany. For cities and regions in a locational competition, it implies an increased need to provide amenities for these high skilled and accordingly high payed workers in addition to a lively firm environment.

Chapter 2

Data and Chapter Overview

The following chapter provides a short overview of the quantitative empirical research design, a chapter overview and a description of the core-dataset.

2.1 Core-Dataset

The empirical analysis of the dissertation relies on a tailor-made dataset that has not been used in the literature before. The dataset consists of geo-coded firm-level panel data for digital firms in Germany.

Such rich, detailed datasets on firms are have rarely been used to date. Especially for German firms, information are hard to retrieve due to data protection. However, with technological progress, this data becomes better available. To date, the 'Mannheim Enterprise Panel' has been used exclusively for German firm data analysis, for example by Pijnenburg and Kholodilin (2014); Audretsch et al. (2015); Fritsch and Aamoucke (2017) and Fossen and Martin (2018). Bersch et al. (2014) provide an overview on the dataset.

The core dataset of this dissertation is provided by North Data (2019). The firm level data originates from statutory publications of German corporations that is the commercial register, commercial register announcements, insolvency announcements and the electronic federal gazette. It encompasses the date of incorporation, date of termination (if applicable), economic field, a description of the company's main business area and an address history. The dataset is not an official dataset, but the data is quasi-official by the virtue of its origin. Technical details in generation are provided by North Data (2019).

The original dataset contains a full export of North Datas' database[1] covering about 3.2

[1]Export 2019 Q2

million German firms in all economic sectors. For the purpose of this dissertation, digital firms have been filtered.

As there is no agreed-upon definition of the digital economy (Duvivier et al., 2018), a digital firm in this dissertation is defined, similar to Weber et al. (2018), as information-technology driven and internet-based. I select firms using NACE codes[2] covering general programming activities, software development, web portals, data processing, and the development of web pages, processing, hosting and related activities and web portals (NACE codes: 62.01.0, 62.01.1, 62.01.9, 62.02.0, 62.03.0, 62.09.0, 63.11.0, 63.12.0.).

Yet, standard industry classification systems have limitations, especially industries that cross over traditional product categories as it is the case for digital firms (Oakey et al., 2001; Bundesagentur für Arbeit, 2019). Since digital business models complement many other sectors, firms may be registered in other NACE codes although running a digital business model. For example, a survey of German startups finds that 31.8 % of new businesses in 2020 were registered in ICT but 66 % in the sample state that they are operating on a digital business model (Kollmann et al., 2020).

Motivated by these shortcomings of NACE code selection, this dissertation takes advantage of the rich information provided in the present dataset. Therefore, the identification of digital firms is broadened by including firms that operate on a digital business model, but formally belong to a different sector. The inclusion of these firms provides a novel approach, which offers a deeper understanding of knowledge flows and the notion of diverse and specific economic inputs in local firm environments (see the discussion of localisation and urbanisation in Chapter 1.1.2). By not including these kind of firms, the dynamics of the digital industry might be biased by a too narrow definition of companies, which benchmarks the main advantages over the 'Mannheim Enterprise Panel'.

In order to identify companies operating on a digital business model, the description of the company's main business area is used. With the help of a word-search selection, firms that are not registered in the ICT sector but operate on a digital business model were added to the dataset. First, the description of the identified ICT firms has been analysed and the most frequently used words related to IT and software have been identified (software development, internet services, IT-services, information technology and programming). Then, these key-words are used to obtain those firms operating on digital business models with the help of several word

[2]The Statistical Classification of Economic Activities in the European Community (NACE) is the classification of economic activities in the European Union.

combinations. Firms that only distribute their products via a web page have been excluded (main key word ”Online Shop”). For the firms that run an online store, key words related to “software development” needed to be included. As an example, a firm that is registered in „Placement of workers“ has been included in the sample, because the objective of the company is “the operation of a social networking platform for skills enhancement and marketing as well as the provision, brokerage and distribution of products and Internet-based services.” Here, Internet-based service has been the selected key-word. Another company registered in ”other livestock farming” develops software for beekeepers and thus their initial knowledge base needs to contain strong digital components.

The resulting sample encompasses firms which are similar in their requirements in terms of employees as well as knowledge; these are the two factors crucial to their competitiveness.

The resulting dataset of digital firms covers a total of 144,230 firms.[3] All firm locations have been geo-coded using the geocode command in R. Each firm has a unique location namely their headquarter, possible subsidiaries are not considered. The location of a firm in a given year is the location as of 31 December. The final panel-dataset consists of firm-year observations starting from the year the individual firms were set up. In the case of firms exiting the market, the firm is deleted from the register and therefore drops out of the data for the year after termination.

One major advantage of the dataset is the precise tracking of locations over time. That is, relocations of firms is traceable on point-level for an exact time. This enables granular capturing of the presence of firms in micro-environments at a given time. Firm relocation is observed based on changes in a unique address identifier in two subsequent years. In general, the data covers a long time span dating back to the 1960s. However, address changes were only tracked digitally after 2007. Thus, when including relocations in the analysis, it is limited to firms born after 2007. For the individual empirical contributions in this dissertation, the data has been aggregated (on distinct spatial scales) and merged with several other datasets, predominantly publicly available administrative data for e.g. population, prices and economic measures such as GDP. Thus, the analysed time periods are determined by the availability of the administrative data rather than firm-data.

The dataset has some limitations. It does not include individual firm information such as financial reports or the number of employees. However, 90 % of all ICT firms have 0 to 9 employees (in 2017), and 7.8 % employ 10 to 49 employees (Destatis, 2021). Therefore I assume

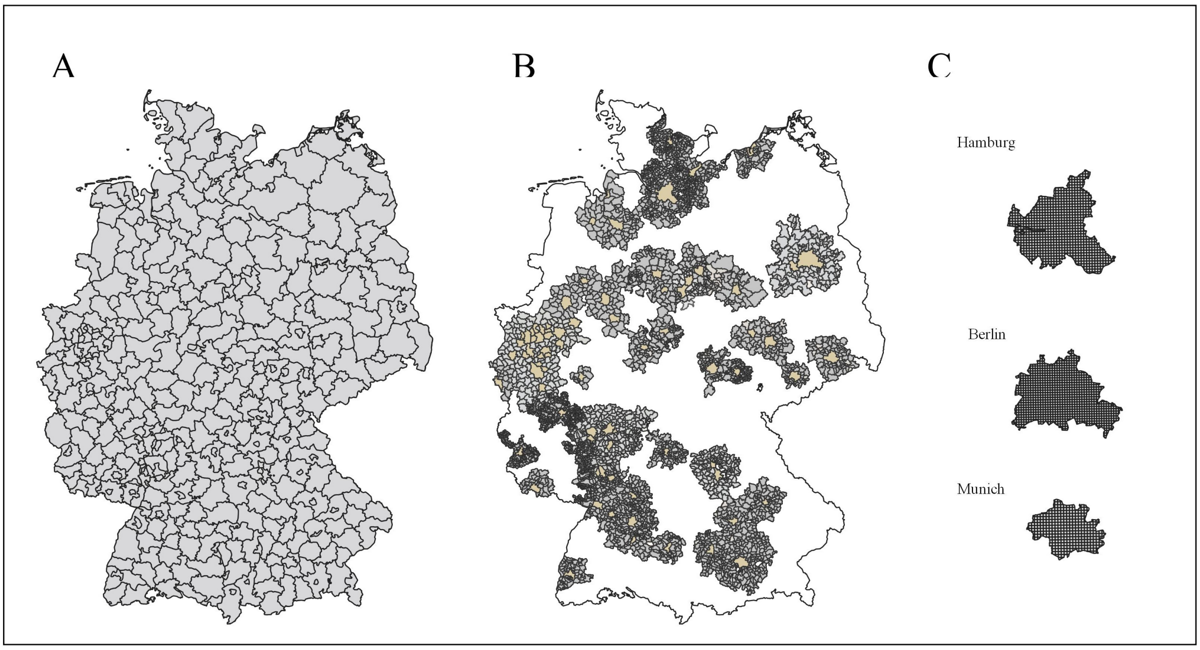
[3]The number of firms is determined by the research setup in the individual chapters and described in detail in the respective data sections.

that the majority of firms in the sample are similarly small in size as it is typical for Germany with its high density of small and medium-sized enterprises (SME)s (Destatis, 2021). Another caveat is that the data only covers one sector, and does not allow an in-depth examination of complementary sectors, especially in terms of co-location, such as venture capital or the role of specific political institutions. When aiming to capture knowledge flows in particular, the location analysis of firms is mainly a proxy. Based on the theoretical backbone of the knowledge spillover theory of entrepreneurship, I assume that profit-maximising firms reveal their preference towards a location that contributes to the firms' competitiveness.

2.2 Chapter Overview

This thesis presents three empirical contributions in Chapters 3, 4 and 5. Each contribution is characterised by its interdisciplinary approach by jointly considering the literature on regional and urban planning as well as economics in a quantitative research design. The contributions share a theoretical basis of literature on agglomeration and commercialisation of new knowledge, i.e. entrepreneurship. Moreover, they present analysis on the common dataset described in the previous section. The different spatial dimensions applied in the individual contributions allow to draw a fine-graded picture of the digital industry within different spatial contexts. As each chapter progresses along the dissertation, the granularity of the spatial analysis increases. Figure 2.1 provides a visualisation of the spatial dimensions. The first contributions' focus is on NUTS 3 regions (Kreise und kreisfreie Städte) (see A), the second contribution focuses on the next level of granularity, that is LAU-Regions (Gemeindeverbände) within labour market regions (B), while the third contribution takes on a micro-approach by investigating location patterns within cities (C).

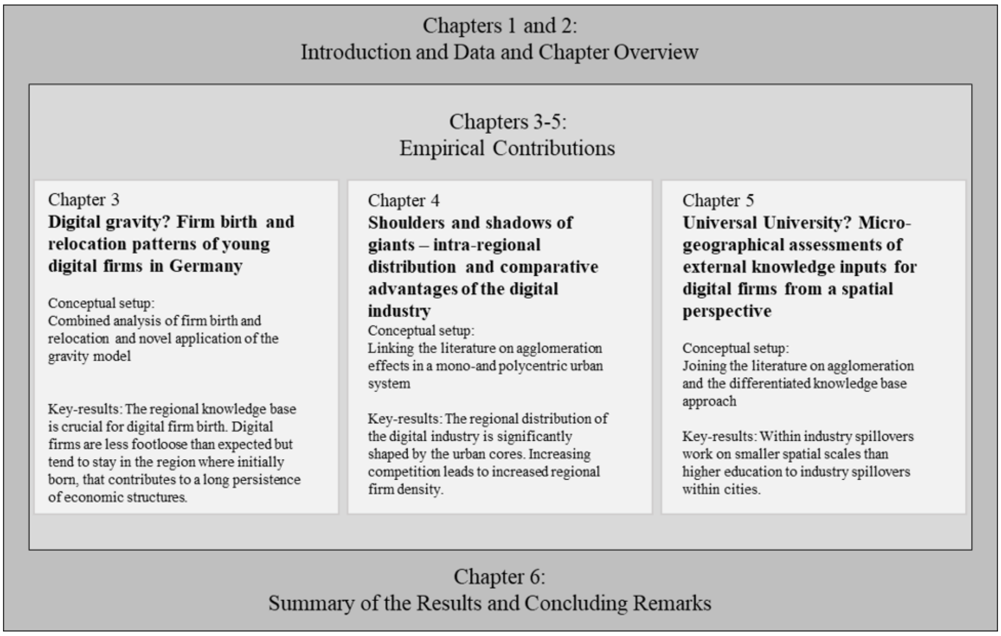
Figure 2.1: Overview on the Spatial Dimensions



The figure gives a schematic overview on the spatial scales employed in this dissertation. Part A displays the county level, Part B shows the municipalities in the urban labour market regions, the core cities are colored in yellow. Note that both maps have the same scale. Part C shows the 1x1 km² level in Hamburg, Berlin and Munich. Datasource basemap: GeoBasis-DE / BKG 2021

Figure 2.2 gives a visualized overview on the individual contributions and main takeaways which will be presented in more detail in the following.

Figure 2.2: Chapter Overview



Source: own illustration.

The first contribution presented in the third chapter analyses and compares firm birth and relocation patterns and its determining factors of the digital economy in Germany between 2008 and 2017 on NUTS 3 level.[4] By considering these two avenues of local industry growth simultaneously, the results draw a detailed picture of the development of the relatively young industry in the long run. Further, it assesses whether digital start-ups and relocating companies favor the same locational characteristics. The out-migration of big industry players of e.g. Oracle, Hewlett-Packard, Tesla, Twitter and Uber from Silicon Valley illustrates the fact that an incumbent location of a firm does not necessarily indicate a long-term impact for the initial location (Duvivier et al., 2018). There are very few papers comparing firm birth and relocation (Holl, 2004; Manjón-Antolín and Arauzo-Carod, 2011; Lee, 2020), while none focuses on the digital industry nor young firms in particular.

On a policy level, the results leave important implications on whether a capitalisation on entrepreneurship capital contributes to a manifestation of the industry in the long run by being

[4]A version of this chapter has been published in the *Journal of Regional Science* and can be accessed via `https://doi.org/10.1111/jors.12624`

able to maintain these new firms. Nevertheless, results show what location factors can contribute to an attractive location for moving firms if own entrepreneurship capital is low.

This chapter uses the core-dataset as described above using the years 2008 to 2016, and thereby a total of 107,321 firms. First, the paper analyses digital firm birth intensity using a linear regression model with Ordinary Least Squares (OLS) with county- and time- fixed effects and a host of regional control variables. The inclusion of the fixed effects allows to control for common trends and captures unobservable county-specific and time-invariant factors which are potentially correlated with the number of new firms. Second, the chapter tests spatial patterns of business relocations building on the theory and operationalisation of the (fixed effects) gravity model. This paper is the first to employ a fixed effects gravity model, the workhorse model for international trade and Foreign Direct Investment (FDI), for the relocation of firms. The model uses aggregated relocation flows instead of individual firm decisions as in discrete choice models, the model most widely used in the relocation literature to date. The advantage of the gravity model is that it includes relocation costs and the geographic dimension. Thus, it enables direct comparison of characteristics in origin and destination in a spatial context. The gravity model captures the spatial dimension by using the physical distance between origin and destination implicitly capturing costs which in turn are increasing in distance. For example, information about a new business site are harder to assess from far away.

Results for digital firm birth show that accessibility of industry-specific knowledge as proxied by the co-location of digital firms is highly conducive to startup activity. For each firm per 1,000 inhabitants, there are 0.18 startups. In other words, almost five digital firms bring about one additional firm birth. In addition, universities are a significant factor for new digital businesses. This result indicates that infant digital firms rely on industry as well as institutional know-how and that locational costs outweigh such locally-bound, tacit benefits, for example knowledge spillovers.

The results of the relocation analysis show significantly more digital businesses relocate to counties with a high density of digital firms. Therefore, digital firms show a strong preference to cluster. I find flows between neighbouring counties are more than twice as big as other flows and relocation flows decay with distance. This is consistent with the fact that moving costs as captured by distance play a crucial deterring role in relocation decisions. This implies a regional persistence of entrepreneurship and that large shifts of the maturing industry are not to be expected.

The main finding from Chapter 3 is that the emergence of the digital industry is a predominantly urban phenomenon with core-periphery dynamics as flows are highest for neighbouring counties. Chapter 4 aims to investigate this interrelations of urban cores and their peripheries more closely. Therefore, it focuses on the spatial distribution of digital firm birth within metropolitan regions. I chose the municipalities (LAU regions) within urban labour market regions as the unit of analysis in order to obtain a more finely graded picture of the developments. This chapter draws on an interdisciplinary approach by investigating industry dynamics in mono- and polycentric urban regions. Papers in the literature on agglomeration effects in urban economics usually assume a standard "core - urban periphery" logic, while possible differences to polycentric urban areas (that is labour markets that host at least two individual urban cores nearby such as the Ruhr-Area) are highly relevant for the planning community. Scholars in the latter mostly focus on population distributions instead of firm behavior. The empirical analysis uses data on digital firm birth between 1995 to 2017 and conducts panel fixed effect regressions OLS for monocentric and polycentric urban labour market regions individually.

Results first show that small municipalities close to core cities gain advantages over their equally small neighbours by hosting a university and from population growth. Second, the regional pattern of the digital industry is shaped by the morphology and digital sector in the closest core cities: Municipalities in monocentric urban regions profit from sharing (population growth) and general knowledge from universities, while municipalities in polycentric urban regions are effected by industry-specific externalities that is an above-average growth in the share of firm birth of their closest urban cores.

The main takeaway from Chapter 3 and Chapter 4 is, that the urban cores significantly shape their regions and the specific knowledge bases being determined by similar firms and knowledge institutions. However, the spatial dimension of the industry dynamics within such urban cores oftentimes remains a black box. Empirical papers on micro-scales within cities are still sparse, mostly due to limited availability of high quality data. This is why Chapter 5 takes on a micro-geographical approach by linking location choices of digital firms within cities to knowledge institutions.

The contribution lies, apart from the spatial unit of analysis, again in the combination of the literature of agglomeration externalities and the differentiated knowledge base approach. The latter has predominantly been applied using qualitative methods to date.

Chapter 5 explores the intra-city knowledge ecosystem in three ways: First, it tests whether there is quantifiable, significant firm clustering close to Higher Education Institution (HEI) and research institutes as this is not entirely clear in the literature.

Second, advantages from localisation (same industry spillover) work on smaller scales than advantages from urbanisation (diverse spillover, e.g. Andersson et al. (2019)). It is almost entirely unclear whether the HEI-industry knowledge exchange requires 'economic proximity', that is similarity of knowledge contents. This is why HEIs are considered separately by their departments and the research area of research institutes. These knowledge contents are then classified into related and unrelated knowledge to ICT firms. Results show whether firm clustering occurs close to any knowledge content or if specific inputs are required.

Third, the HEI landscape in Germany consists of research universities, University of Applied Science (UAS) (technical colleges) and universities of music, arts and design (hereafter design universities) that differ substantially in their institutional setups. Based on differentiated knowledge bases, each type of HEI is assigned one knowledge base. Research universities inhibit highly codified analytical knowledge that can be transferred over long distances. UAS rather host more engineering-based, problem solving, applied synthetic knowledge that is not as easily transferred over long distances as analytical knowledge. Finally, symbolic, highly context specific knowledge requires face-to-face contacts when transmitted and is mostly found in design universities. I test whether firm clustering occurs in accordance to the knowledge bases' transferability on micro-scales leaving important takeaways for urban policies.

Results show that firms cluster significantly in neighbourhoods with HEIs and research institutes, but cluster effects within the industry decay more rapidly over distance than for institutional knowledge sources. Thus, tacit spillovers from knowledge institutions are conducive to new firm births. Further, university-industry spillovers benefit many firms within the city and are not limited to very few firms that select into the 'right' neighbourhoods.

There is significantly more firm birth in neighbourhoods offering specialised institutional knowledge on IT and data science, research institutes for social sciences and design departments. This is reflective of a digital firm's desire to be closest to a knowledge stock that can directly be transferred into products.

Additionally, there is no significant effect for firm birth close to research universities, while there are strong positive effects for UAS and design universities. This novel finding is of great relevance for policy makers as universities in particular are an often-used outlet for public funding to accelerate firm birth activity. In Germany, this is often done by subsidizing office spaces close to (research) universities. This fifth chapter shows that a fine-graded political funding and possibly an industry-specific location policy vis-á-vis co-location to a HEI are needed.

Finally, Chapter 6 provides a summary of the main findings, conclusions and policy implications with respect to the empirical findings as laid out in the individual chapters.

Chapter 3

Digital Gravity? Firm Birth and Relocation Patterns of Young Digital Firms in Germany

Abstract[1]

This Chapter analyses the spatial patterns of young (<10 years) digital firms in Germany from 2008 to 2017 on county level. Determinants of firm birth locations and relocations are considered jointly to understand differences in location choices within firms' life cycles. I match commercial register data of 107,321 firms with county level administrative data to capture local characteristics. Using an OLS model, I find the local knowledge base – universities and co-located incumbents – to be a significant key determinant of digital firm birth when controlling for a host of local characteristics. My results indicate that for five firms per 1,000 inhabitants, there is around one firm birth. Second, using a fixed effects gravity model for the analysis of relocations, I find that the most dominant explanatory factor for firm relocation across specifications is distance, that is, relocation costs. Relocation flows are more than twice as high to neighbouring counties relative to other locations showing that digital firms are not as footloose as their business model may suggest. Jointly, my results reflect economic activity's regional persistence, particularly for new firms. This Chapter provides evidence for policies targeting homogeneous digital clusters based on strong co-location. Digital economic activity is not shifted over long distances, but the regional entrepreneurship capital is crucial for local growth.

[1]A version of this chapter has been published in the *Journal of Regional Science* and can be accessed via `https://doi.org/10.1111/jors.12624`

3.1 Introduction

Scholars observe a recent trend of young, high-tech and knowledge-intensive firms to locate in clusters in central districts and neighbourhoods in large cities (see e.g. Foord (2013); Duvivier et al. (2018)), seemingly challenging the Silicon Valleys of the Western World. Part of this trend is not only driven by newly found startups but also by firms relocating to cities. For example, Twitter, Uber and Airbnb have chosen to set up their new headquarters in downtown San Francisco, while Oracle, HP and Tesla chose to relocate to 'Silicon Hills' in Austin, Texas (Duvivier et al., 2018). As the big players in the technology industry bring employment, taxes and a host of other externalities for the region, their relocations exemplify just how important location decisions of both new and existing firms are for regional economic development (Audretsch et al., 2006; Fritsch and Mueller, 2008).

Thus, there are a range of public efforts often using strong subsidisation to support and foster digital economic activity by means of attracting new businesses, incubating startups, developing and/or supporting technology hubs and networks. To this end, understanding the factors which determine locational choices for start-ups and incumbents' relocations are crucial for efficient policy making (Lee, 2008). Initial location choices and the role of regional factors for entrepreneurship are well studied and the literature shows that start-up hot-spots are highly persistent over time (Fossen and Martin, 2018). At the same time, young knowledge-intensive firms become even more likely to change locations (Esteve-Pérez et al., 2018) in order to grow and gain competitiveness (Stam, 2007; Guzman, 2019).

This Chapter uses commercial register data on 107,321 young digital firms in German counties (NUTS 3) between 2008 and 2017 to analyse firm birth and relocation patterns in the high-tech industry. The main objective is to compare regional preferences in firm births and relocation patterns to assess whether digital start-ups and relocating companies favor the same locational characteristics. If they do not share the same location requirements, more tailored policies towards those two avenues of local economic growth are necessary.

First, I analyse digital firm birth intensity using a linear regression model (OLS) with county- and time- fixed effects and a host of regional control variables. Results show that accessibility of industry-specific knowledge as proxied by the co-location of digital firms is highly conducive to startup activity. For each firm per 1,000 inhabitants, there are 0.18 startups. In addition and in line with the literature, universities are a significant factor for new digital businesses. This result indicates that infant digital firms rely on industry as well as institutional know-how and that locational costs outweigh such locally-bound, tacit benefits.

Second, I test spatial patterns of business relocations building on the theory and operationalisation of the (fixed effects) gravity model. Counties with higher agglomeration benefits such as a specialized high-tech labour market and the potential for IT-specific knowledge spillovers are expected to attract more relocating firms. Additionally, most relocations should occur between geographically proximate or contiguous counties where moving costs are low while access to locally bound factors such as local customers, suppliers or networks remains relatively low-cost.

The results of this empirical exercise show that significantly more digital businesses relocate to counties with a high density of digital firms. Therefore, digital firms show a strong preference to cluster. I find flows between neighbouring counties are more than twice as big as other flows and relocation flows decay with distance. This is consistent with the fact that moving costs as captured by distance play a crucial, deterring role in relocation decisions. These findings are at odds with the common perception that digital businesses are relatively footloose (Weterings and Knoben, 2013; Esteve-Pérez et al., 2018). On the contrary, my findings indicate that they tend to stay in their regions of origin. This finding, however, is in line with Knoben (2011), that firms with high dependency on outside resources and strong networks do not relocate over long distances, indicating industries' regional persistence (Fritsch and Storey, 2014). Moreover, the joint investigation of firm birth as well as incumbents' relocation patterns reveals that policies targeting digital firm birth also spill over into neighbouring counties in the medium to long run next to the expected initial local benefits. Therefore, an intra-regional cooperation strategy to foster firm birth where counties administrations pool their resources would be a promising approach.

This Chapter contributes to the literature in three ways. First, very few papers allow direct comparison of firm birth locations and relocation (Holl, 2004; Stam, 2007; Manjón-Antolín and Arauzo-Carod, 2011; Lee, 2020) while none – to the best of my knowledge – focuses on the relocation of "infant" firms. Nonetheless, a joint analysis of new and young digital firms is highly informative in order to understand the development of regional patterns of entrepreneurship in the long run. Moreover, it is new as well as young digital businesses which drive industry growth next to the long-established big players in the industry. Especially young, growing firms are a potential asset for a county's economic development. Thus, understanding this particular neck of the digital industry and its spatial preferences matters greatly in the context of political interventions aimed at alleviating regional disparities and smoothening structural economic transitions by means of financially supporting the digital industry.

Second, many of the existing papers on relocation focus on manufacturing (Holl, 2004; Conroy et al., 2016; Yi, 2018) or compare different sets of industries along the business models' knowledge-intensities (Kronenberg, 2013; Nguyen et al., 2013; Weterings and Knoben, 2013). Yet, digital business models similarly to other service sector industries have a different cost-structure than manufacturing – labour intensive while sunk costs are low – thus requiring separate consideration. Moreover, digital companies provide broadly applicable technology which affords productivity gains to almost all other sectors and therefore differ from other industries in terms of market reach and locational choices.[2]

Third, I make a methodological contribution to the literature by applying a state-of-the-art gravity model framework to the analysis of firm mobility. This is in contrast to the majority of regional studies that rely on discrete choice models which neither include relocation costs nor mobility's geographic dimension. Moreover, these specifications are rather ad-hoc without theoretical priors on the key decision factors and mechanisms driving the results. In the literature on FDI, however, the gravity model has been (theoretically) established (Portes and Rey, 2005). Since FDIs are essentially cross-border firm- or subsidiary relocations, there is reason to believe that the gravity model is appropriate for modeling firm relocation flows.

As the workhorse model in the migration and international trade literature, the gravity model captures the spatial dimension by using the physical distance between origin and destination implicitly capturing costs which in turn are increasing with distance. For example, information about a new business site are harder to assess from far away. The key advantage of using the gravity model with a full set of origin- and destination- fixed effects for the analysis of firm mobility is that it allows for modeling one of the key unobservables, namely implicit and explicit relocation costs, while also controlling for location-specific characteristics and common time trends affecting all locations. The inclusion of these fixed effects is important for identification as a lot of the variation in the data can be explained by these location-specific characteristics. Thus, the remaining significant determinants in the model can be interpreted beyond any time and location-specific confounding factors.

The remainder of this Chapter proceeds as follows. Section 3.2 contextualizes the analysis in light of the existing literature while Section 3.3 embeds it in a theoretical framework. Section

[2]While there are many studies on entrepreneurship (e.g Audretsch et al. (2006); Pijnenburg and Kholodilin (2014); Fossen and Martin (2018)) in Germany, there are hardly any studies about firm relocation.

3.4 describes the data in detail and Section 3.5 the empirical strategy. In Section 3.6, I present the findings: first my findings on firm birth patterns, then the results for the gravity models of ICT firms' relocations in Germany. The last section concludes, while also highlighting the key takeaways for policy makers and regional planners.

3.2 Literature on Regional Determinants of Firm Birth and Relocation

3.2.1 Firm birth

Factors which are conducive to entrepreneurship as measured by high local startup rates have been well studied in the literature (Glaeser et al., 2010; Fritsch and Storey, 2014; Buczkowska and de Lapparent, 2014; Glaeser et al., 2015). In a study of West-German new technology-based firms, Bade and Nerlinger (2000) find the highest startup rates relative to the labour force in close proximity to core cities, while core cities have the most startups in absolute terms. Van Oort and Atzema (2004) find ICT firms to co-locate in areas with dense economic activities. Moreover, Audretsch et al. (2012) find that local employees' propensity to start a business is highest in urban agglomerations and their periphery. In this context, Pijnenburg and Kholodilin (2014) link entrepreneurship capital and knowledge-based startup rates. They find that knowledge spills over from its source to the startup also across NUTS 3 borders. Since urban agglomerations innately offer a diverse economic environment they provide key factors to thrive for digital businesses. I thus expect high startup rates in core cities and their surroundings.

3.2.2 Relocation

The focus in this chapter lies on location patterns of new digital businesses and firms younger than ten years. The probability of relocation is very high within the first ten years of a firm's life: In order to reduce risk of failure, firms have to innovate and reconsider products, activities and eventually their location (Esteve-Pérez et al., 2018; Rossi and Dej, 2020). Apart from individual firm characteristics, a firm's decision to relocate can be motivated by external location-specific characteristics that 'push' the firm away from its current location such as steeply rising real estate prices (Van Dijk and Pellenbarg, 2000; Van Wissen, 2000). Relocating firms can be 'pulled' into regions offering more suitable location characteristics (Holl, 2004; Kronenberg, 2013). Agglomeration benefits are strongly identified as pull factors since firms across all sectors

- in particular services - are drawn to densely populated municipalities to benefit from higher local demand, stronger, better educated workforces and a wider supply of local public amenities (Bodenmann and Axhausen, 2012; Weterings and Knoben, 2013; Kronenberg, 2013; Nguyen et al., 2013; Rossi and Dej, 2020).

Typically, service sector firms want to benefit from the locally bound knowledge as well as the labour pool. Thus, they are attracted to dense, high-quality-of-life municipalities in spite of their high sector-specific wages (Kronenberg, 2013). That is firms do not necessarily adopt a pure cost minimisation strategy but choose locations where they can be certain that high-skilled workers and necessary amenities are available (Rossi and Dej, 2020). Stam (2007) argues that relationships with social networks are especially important in the early stages of a firm's life while cost considerations become more important later, while Knoben (2011) finds firms being dependent on outside resources tend to move short distances. Presumably, knowledge-intensive digital firms do not need as much space as manufacturing firms with constant access to knowledge being more important than a low-cost location.

3.3 Theoretical Framework

3.3.1 Agglomeration effects and firms' location decisions

Firms derive advantages from agglomeration externalities, especially in cities. The main benefits are the pooled labour market, access to specialized suppliers and benefits from knowledge spillovers (Krugman, 1991; Armington and Acs, 2002). These agglomeration effects can be divided into localisation and urbanisation economies which are considered in turn.

Urbanisation effects refer to the benefits of diversity and density of amenities such as public infrastructure which cities typically offer (Jacobs, 1969). Next to the size of the labour market, these are in particular the urban density of universities, research institutes and other knowledge- and research-related activities facilitating knowledge spillovers between firms. Novel knowledge and innovation is closely linked to entrepreneurship through commercialisation of knowledge into new firms (Acs et al., 2009). Thus, a vital regional knowledge base is more likely to be bigger in cities rather than regional agglomerations rendering cities particularly attractive locations for young firms.

Localisation describes co-location of similar firms in close proximity (Porter, 1990) with firms mainly benefiting from employment advantages such as specialized employees and a lower probability of labour shortages (Krugman, 1991). Particularly important benefits for software firms are knowledge spillovers accruing from inter-firm cooperation as well as fluctuation of employees (Trippl et al., 2009). Moreover, co-location of related industries fosters entrepreneurship by lowering costs of starting a business for individuals and enabling better access to a more diverse range of inputs and complementary goods (Glaeser and Kerr, 2009).

Novel knowledge and innovations are key for profits and growth and thus are particularly relevant for location analysis. General innovations are typically developed and harbored in universities and research institutes while incumbent firms typically hold an advantage in new marketable products. Now, if the expected value of a certain piece of knowledge is higher for an individual than for the decision maker in the institution or the firm this individual will start a new business if costs are low (Acs and Varga, 2005; Acs et al., 2009). There is an interdependence between existing knowledge institutions and industry players which is conducive to entrepreneurship within the spatial reach of such knowledge spillovers (Bade and Nerlinger, 2000; Trippl et al., 2009; Fritsch and Aamoucke, 2013). Furthermore, university graduates are a source of qualified labour supply to local firms. This can be advantageous for relocated firms if the labour market at their origin is insufficient (Armington and Acs, 2002).

3.3.2 Firm birth and mobility

The fundamental difference between location and relocation theory is that relocations explicitly substitute one location for another, while newly established firms are not constrained by previous location decisions. In general, a firm moves from its current location if the location is no longer inside the spatial margins of profitability (Brouwer et al., 2004; Ozmen-Ertekin et al., 2007). Relocation decisions are likely to be explained by the differences between origin and destination in terms of profitability as well as relocation costs.

This chapter moves away from analysing individual firm movements and relies on an aggregate approach distinguishing between inter- and intra-regional migrations. Businesses intra-regional moves amount to industrial sub-urbanisation around larger urban agglomerations (van Dijk and Pellenbarg, 2017). This is referred to as the incubator hypothesis which postulates that manufacturing firms are born in central urban areas and they out-migrate to urban peripheries in their growth phase in order to find expansion space at a location that is easily accessible for clients and suppliers (Leone and Struyk, 1976; van Dijk and Pellenbarg, 2017).

Entrepreneurs tend to disproportionally take their hometown as a natural firm birth location and thus exhibit a strong home bias (Figueiredo et al., 2002; Michelacci and Silva, 2007). This local entrepreneurship capital is an element of the region's endogenous economic potential (Stam, 2007; Fritsch and Wyrwich, 2014). Conceivably, home-biased entrepreneurs located in the periphery revise their initial location decision once they have proven viable and move to nearby cities in order to benefit from agglomeration advantages. Inter-regional moves involve industrial decentralisation from economic core areas to peripheral areas. Within this type of movement, firms move to areas with lower land and/or labour costs (van Dijk and Pellenbarg, 2017). The increase in land prices induced by agglomeration of economic activities could be a more significant incentive for older, established firms to opt-out of industrial agglomerates as the cost-benefit trade-off becomes unfavorable for the urban location (Combes et al., 2012).

Moves between core cities often reflect firms moving from diversified (urbanisation economies) into specialized cities (localisation economies) (Duranton and Puga, 2001). Businesses start up in diversified cities (urbanisation economics) until they find an ideal business process and ultimately relocate into a specialized city (localisation economics) when switching to mass production (for manufacturing firms). Systematic comparisons of location patterns of start-ups and relocating firms such as by Holl (2004) and Manjón-Antolín and Arauzo-Carod (2011), for example, indicate that startup activities are highly associated with industrial diversity while firm relocations are not.

3.3.3 Theoretical Predictions and Hypothesis

The conceptual framework of this Chapter draws on the theoretical and empirical findings on spatial patterns of firm birth as laid out above. Agglomeration benefits stand out as a crucial location factor for firm birth and entrepreneurial activity. I thus expect young digital firms to display a strong preference towards cities (Hypothesis 1).

With regard to localisation and urbanisation, I use service - and industrial ratios to capture the general local economy and its specialisation. With increasing co-location of similar firms (digital firms per 1,000 inhabitants), theoretical agglomeration mechanisms predict increasing firm birth rates. As local rent-prices display a firm's willingness to pay for agglomeration benefits, they are included in the empirical analysis. As a knowledge-intensive industry, firms in the digital economy are dependent on several knowledge providers such as universities and research institutes and their possible labour markets. Universities and research institutes are expected to have a positive effect on firm births (Hypothesis 2).

Applying the gravity model to digital firms' relocation patterns, I expect that agglomeration benefits vary considerably with firm age as well as other firm characteristics and thus individual firms may find agglomerates attractive according to their individual preferences as captured by gravitational force of the associated agglomeration benefits. One such example could be a maturing firm seeking to innovate or improve their products or to broker into a new market. Therefore, I hypothesize that agglomerated areas receive higher in-flows (Hypothesis 3).

Digital firms do not require a lot of physical space neither when starting out nor when growing since scaling up of digital products has different space requirements than manufacturing - the most common sector of investigation in the relocation literature.[3] Therefore, I do not expect firms to select out into peripheral areas due to an unfavorable trade-off between costs and access to knowledge and market, but rather to remain within the agglomeration effects' spatial margins. Thus, consistent with the assumptions of the gravity model, relocation flows are expected to decrease with distance (Hypothesis 4).

3.4 Data

The tailor-made dataset encompasses a panel of 107,321 digital companies in Germany that stands out for its precise tracking of firms' locations over their lifecycles (see Section 3.4.1) down to the address level. This firm level data are combined with regional characteristics from several sources with a strong focus on those which are relevant for digital businesses and the knowledge economy more generally.

3.4.1 Firm Data

The core dataset is provided by North Data (2019). The firm level data originates from statutory publications of German corporations.[4] It encompasses the date of incorporation, date of termination (if applicable), economic field, a description of the company's main business area and address history. Since address changes were only tracked digitally after 2007, the analysis covers companies that have entered the market between 2008 and 2017.

This data does not include individual firm information such as financials or the number of employees. However, 90 % of all ICT firms have 0 to 9 employees (in 2017), and 7.8 % employ 10

[3] Moreover, it is conceivable that digital firms compete through the price channel rather based on scale as is typically the case for service industries (see e.g. Saarenketo et al. (2008)).

[4] Commercial register, commercial register announcements, insolvency announcements and electronic federal gazette. The dataset is not an official dataset, but the data is quasi-official by the virtue of its origin. For details in data generation see North Data (2019).

to 49 employees (Destatis, 2021). Therefore I assume that the majority of firms in the sample are similarly small in size as is typical for Germany with its high density of SMEs (Destatis, 2021).

As there is no agreed-upon definition of the digital economy (Duvivier et al., 2018) for the purpose of this Chapter, a digital firm is defined as information-technology driven and internet-based. I select firms using NACE codes similar to Weber et al. (2018) covering general programming activities, software development, web portals, data processing, and the development of web pages, processing, hosting and related activities and web portals.[5] Yet, standard industry classification systems have limitations, especially industries that cross over traditional product categories as is the case for digital firms (Oakey et al., 2001).

Since digital business models complement many other sectors, firms may be registered in other NACE codes although running a digital business model. For example, a survey of German startups finds that 31.8 % of new businesses in 2020 were registered in ICT but 66 % in the sample state that they are operating on a digital business model (Kollmann et al., 2020). To tackle this issue the description of the company's main business area is used. With the help of a word-search selection, firms that are not registered in the ICT sector but operate on a digital business model were added to the dataset.[6] The resulting sample encompasses firms which are similar in their requirements in terms of employees as well as knowledge; that is, in terms of the two factors crucial to their competitiveness. In total, 107,321 firms are covered in the dataset over the sample period of nine years.

The location of a firm in a given year is the location as of 31 December. Each firm has a unique location namely their headquarter, possible subsidiaries are not considered. The panel-dataset consists of firm-year observations. Firm relocation is observed based on changes in the 5 to 6 - digit unique county identifier in the two subsequent years. Firms which exit the market are deleted from the firm register and excluded from the panel after the year of deletion. In

[5] 62.01.0, 62.01.1, 62.01.9, 62.02.0, 62.03.0, 62.09.0, 63.11.0, 63.12.0.

[6] First, the description of the identified ICT firms has been analysed and the most frequently used words related to IT and software has been identified (Software Development, Internet services, IT-services, information technology and programming). Then, these keywords are used to obtain those firms operating on digital business models with the help of several word combinations. Further, firms that only distribute their products via a webpage have been excluded (Keyword Online Shop). For those firms, key words related to "software development" needed to be included. As an example a firm that is registered in „ Placement of workers“ has been included in the sample, because the objective of the company is "the operation of a social networking platform for skills enhancement and marketing as well as the provision, brokerage and distribution of products and Internet-based services." Here, Internet-based service has been the selected key-word.

total, 12.27 % of the resulting sample relocated to a different county between 2008 and 2017 amounting to 14,878 moves.[7]

The location is available on point-level but is aggregated to the county level (NUTS 3). The analysis of a panel of firms on NUTS 3 level thus moves away from observing individual firms, which provides insights into regional dynamics. Moreover, meaningful policy implications can be drawn based on the same spatial unit of analysis as relevant for policy makers seeking to foster entrepreneurship.

3.4.2 Regional Characteristics

The aggregated data is merged with several datasets containing regional characteristics for all 401 German counties between 2008 and 2017. The majority of the data is retrieved from the INKAR database published by the German Institute for Research on Building, Urban Affairs and Spatial Development (Bundesinstitut für Bau-, Stadt- und Raumforschung (BBSR)).[8]

Population density (BBSR, 2020) is used as a measure of local market potential as well as urbanisation economies and agglomeration benefits (Rosenthal and Strange, 2008; Mameli et al., 2014). As a proxy for local price levels, I use an apartment rent real-estate price index originating from Immobilienscout24, the largest German real-estate search site, which is provided by the Leibniz-Institut für Wirtschaftsforschung (RWI) (Klick and Schaffner, 2021). The RWI's index captures the difference of the counties' mean rent price to the German average price.[9] Housing costs reflect the willingness to pay for agglomeration benefits (Combes et al., 2019) and the price differential proxies high-cost locations, in particular, which can be assumed to offer the highest benefits in line with costs. Lastly, to measure labour costs, the local average gross income is used (BBSR, 2020).

Capturing specialisation and localisation mechanisms, an industrial and a services ratio are used. Both capture the percentage of employees per 100 inhabitants of working age in the respective sector (BBSR, 2020). Moreover, to measure regional co-location of same-sector firms (localisation), the number of digital firms per 1,000 inhabitants that entered the market after 2007 in the respective county is included. This density measure of incumbent firms also is indicative of industry depth as well as regional specialisation.

Locally available knowledge and research intensity are proxied by the number of universities

[7]In total there are 1,571 firms that move more than once during the sample period.
[8]For an overview of data on regional characteristics and sources see Appendix Table A1.1
[9]For more information see Klick and Schaffner (2021).

and publicly funded research institutions in the respective counties. To this end, the locations of research institutes belonging to the four major German research associations (Fraunhofer Institut, 2019; Helmholz Gesellschaft, 2019; Leibniz Association, 2019; Max-Planck-Institute, 2019), major research institutes funded by federal states as well as the national government (Forschungseinrichtungen des Bundes und der Länder, (OEFW, 2016)) are included. For universities, locations published in a public register of colleges and universities are used (Hochschulkompass, 2020).[10] Both, research institute and university data are time-variate although the variation is not large with only 31 new colleges which amounts to 6 %. Table A1.1 in the Appendix provides an overview on the regional data.

3.4.3 Descriptive Statistics

Firm Birth

In Table 3.1, summary statistics for startups and county level regional characteristics for the period from 2008 to 2017 are presented.

Table 3.1: Summary Statistics

Statistics	N	Mean	St. Dev.	Min	Max
Popular density	4,010	522.511	681.476	36	4,713
Industrial ratio	4,010	17.934	8,621	4.900	91.800
Service ratio	4,010	35.548	14.608	12.900	96.600
Gross income	4,010	2.381	0,381	1.598	4.367
firm birth / year	4,010	24.926	79.199	0	1,482
Firms per 1,000 inhabitants	4,010	0.507	0.439	0.009	5.283
Price index	4,010	-10.616	18.151	-84	119
Price index change	4,010	6.571	8.644	-13	85
Research institutes	4,010	1.402	3.826	0	45
Universities	4,010	1.246	2.517	0	33

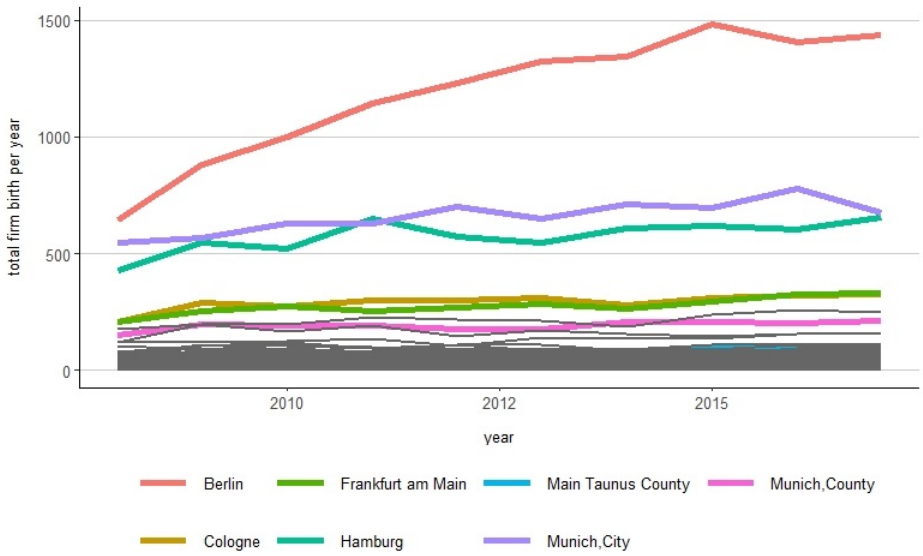
Notes: The table shows summary statistics for firm birth on county level and regional characteristics. The sample includes 401 German countries and covers all years from 2008 to 2017. See Table A1.1 for more information on data sources and definitions.

There are 25 new businesses each year in the average county, while Berlin registered 1,482 new digital firms at its peak in 2015. Figure 3.1 presents the regional distribution of total firm births.

In absolute terms, Berlin is the top location for digital firms. Hamburg and Munich rank

[10] Firm addresses have been geocoded with the geocode command in *R*.

Figure 3.1: Total Firm Birth of Digital Firms in Counties 2008 to 2017



Notes: The Figure shows the total firm birth of digital firms in counties 2008-2017. The top seven counties are colored as displayed in the legend. All other counties in grey.

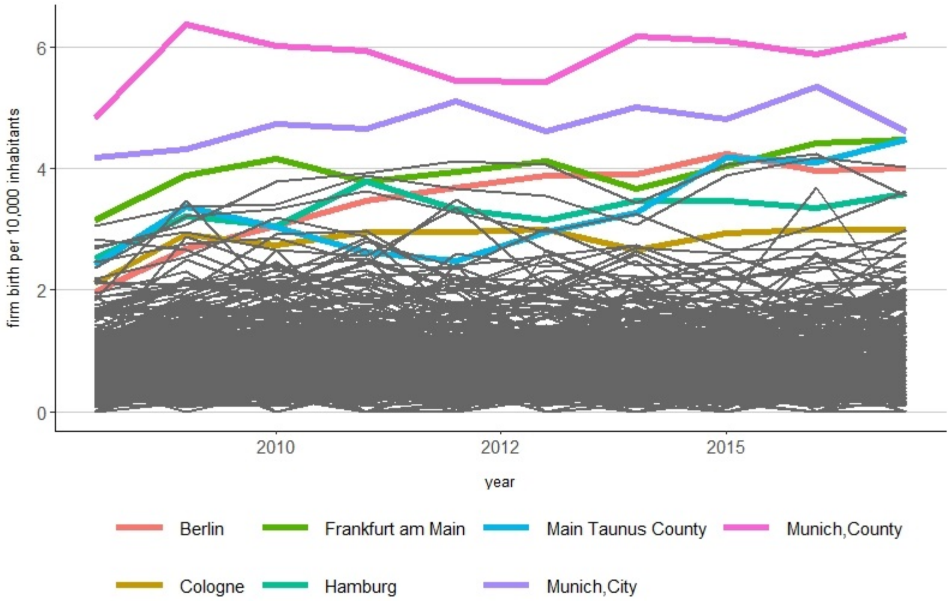
second and are almost on par with at 655 and 674 new digital firms respectively in 2017, followed by Cologne (324 firms in 2017) and Frankfurt/Main (325 firms in 2017). It is worth noting that the top six locations for firm births over time are cities with more than 500,000 inhabitants underpinning the fact that start-up culture is most pronounced in urban areas.

Figure 3.2 presents firm births relative to the population. Munich County – neighbouring Munich City – has the highest firm birth per 1,000 inhabitants. Also, Main-Taunus-County – neighbouring Frankfurt/Main – shows high firm births per 1,000 inhabitants. The descriptive statistics indicate – consistent with theory and similar to Bade and Nerlinger (2000) findings – that absolute startup rates are highest in core cities, while firm birth relative to the population can also be high in the core areas' periphery.

Firm mobility

There are 10,108 out of 1,443,600 origin-destination county pairs with positive relocation counts over the study period (see Table 3.2), that is 0.7 % of observations. When pooled over the whole sample period, there are 3.6 % positive flows. Compared to the 12 % of firms moving to different counties, this aggregated flow appears to be very small. The reason for this is that the data is highly spatially dispersed as 80 % of the origin-destination pairs only have one movement in a given year, accounting for 54.4 % of all flows. That means that 20 % of the positive observations

Figure 3.2: Firm Birth of Digital Firms per 10,000 inhabitants in Counties 2008 to 2017



Notes: The Figure shows the firm birth of digital firms per 10,000 inhabitants in counties 2007-2017. The top seven counties are colored as displayed in the legend. All other counties in grey. Datasource basemap: GeoBasis-DE / BKG 2021

cover 45 % of all relocations. Strikingly, the largest single relocation count is 58 businesses moving from Munich City into Munich County in 2017 firms (29.9 % of businesses out-migrating Munich City). To contrast this, there is only one county (Hildburghausen) which has never been a destination of a relocating digital firm within the sample period.

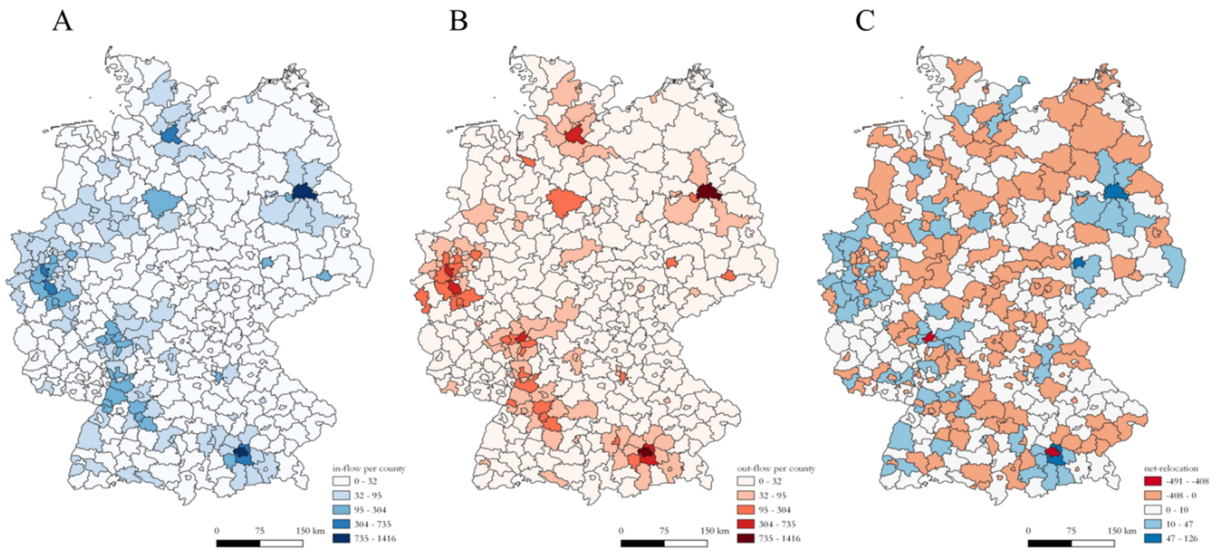
Table 3.2: Descriptive Statistics for Bilateral Flows and Distance

Statistic	N	Mean	St. Dev.	Min	Max
Bilateral flow	1,443,600	0.01	0.20	0	58
Distance in km	1,443,600	304.40	151.66	1.26	824.48
Bilateral flow only positive	10,108	1.47	1.85	1	58

Notes: Sample covers the year 2008 to 2017.

When looking at counties receiving firm inflows, movements between the City of Munich and County Munich and vice versa accounts for the largest relocation flows with 539 firms (3.6 %) relocating in total. The third and fourth largest flows occur from firms migrating from Munich to Berlin and from Hamburg to Berlin, respectively. Table 3.2 presents summary statistics on the bilateral flows and distance. Figure 3.3 shows relocation activity geographically, where A shows the absolute inflow of firms in a distinct county for all years, B shows out-flows per county and C the net-relocation.

Figure 3.3: Firm Migration Patterns



Notes: A displays the inflows of firms per county, B the out-flows per county and C the net-relocation. All flows have been summed up to all examined years. Basemap: GeoBasis-DE / BKG 2021.

In C, the counties marked in red show a negative net-relocation, while blue colored counties have an overall surplus of in-migrating firms. We see that large cities like Munich, Hamburg and Berlin have both strong in- and outflows. However, Berlin is the only metropolitan city gaining a surplus of digital firms while counties in the surroundings of big cities like Munich and Frankfurt/Main seem to attract relocating firms.

In sum, the figures reveal two relocation patterns. First, it shows firm migration flows into big cities, in particular big inflows to Berlin, which is indicates that digital firms behave in accordance with Duranton and Puga (2001) model; that is, they move to a specialized city with the highest absolute firm birth rates as is the case for Berlin. Second, we see a striking pattern of an urban-core and urban-periphery dynamic.

Moreover, the strong persistence in the data is reflective for a very persistent relocation behavior of firms with only few counties being deemed attractive for relocating firms. As much as this pattern is in accordance with expectations it also poses econometric challenges. In terms of the method used, the gravity model is geared particularly well to such persistence in observations and spatial disparities. In terms of the research question, it provides a framework for understanding the factors explaining these high flows into particular counties. The empirical- as well as the robustness-section (Chapter 3.7) will deal with this empirical issue in detail.

3.5 Empirical Strategy

Empirically, three main models are estimated. First, I model the initial location of young digital firms on county level using annual data covering the years 2008 to 2017 as a function of regional characteristics of the firms' birth locations (Section 3.5.1). Second, to model relocation, I use the same set of variables in a gravity model which considers both origin and destination counties, in two specifications (Models 2 and 3, as presented in Section 3.5.2).

3.5.1 Firm birth analysis

To explore the initial location choice of new digital businesses, two pooled and fixed effect OLS models are employed respectively. Model 1.1 and 1.3 are fixed effects models. Model 1.2 and 1.4 are pooled models (without fixed effects) mainly to benchmark against the literature. The baseline model is will be estimated in several versions presented in (Equation 3.1):

$$\ln(firmbirth_{i,t}) = \alpha + \ln(l_{i,t}) + G_i + NG_i + T_t + \gamma_i + \epsilon_{i,t.} \tag{3.1}$$

The dependent variable $\ln(firmbirth_{i,t})$ is digital firm birth per 1,000 inhabitants in location i at time t for Model 1.1. In Model 1.3, the dependent variable is the absolute firm birth in location i at time t. The independent variables enter all models as follows: $\ln(l_{i,t})$ refers to a vector of of locational characteristics in location i at time t (population densities, service ratio, industrial ratio, gross income, price index (change), universities, research institutes).

All location characteristics enter the model as laid out in the data section above. G is an indicator variable equal to 1 if the county is a city with more than 500,000 inhabitants, zero otherwise. NG is an indicator variable equal to 1 if the counties share a border with a city of more than 500,000 inhabitants capturing spatial externalities of agglomerations (zero otherwise). $\epsilon_{i,t}$ is the error term. For models 1.1 and 1.3, T_t is a year fixed effect controlling for common trends, γ_i is the time-invariant county fixed effect capturing unobservable county-specific and time-invariant factors which are potentially correlated with the number of new firms. These do not enter the pooled models 1.2 and 1.4.

3.5.2 Models for the analysis of firm relocations – Gravity model

Firm relocation is typically modelled by individual firm location decisions following the choice model of McFadden (1973) (Arauzo-Carod et al., 2010; Bodenmann and Axhausen, 2012; Kronenberg, 2013). Here, the profit-maximizing, fully-informed firm decides to move after having screened profitability prospects in all possible alternative locations. The estimation outcome captures the probability to move to a particular location given the destination's characteristics.

A more precise understanding of the factors motivating relocations, however, requires modeling firms' movements between locations, that is, a direct differential between origin and destination characteristics to explain firms' bilateral flows (Kohler, 1997; Conroy et al., 2016). The key underlying assumption is that the new location reflects the firm's 'revealed preference' and thus is the result of a curating process of all relevant alternatives. A standard theoretically embedded model for bilateral flows in a monopolistic competition environment, as it should be the case for the digital economy, is the gravity model.

The gravity model is the workhorse theoretical framework in empirical analysis of bilateral flows in international trade as well as their spatial determinants (Head and Mayer, 2014; Yotov et al., 2016). However, it has been applied in many other fields, such as labour migration (Karemera et al., 2000), tourism flows (Morley et al., 2014), or the selection of airline hubs as an example of industry location (Drezner and Drezner, 2001; Redding et al., 2011), also see (Head and Mayer, 2014, p. 148f.). The key similarity between these literature's and this Chapter is that production factors and/or facilities are relocated across spatial units.

Closest to the line of analysis in this Chapter is the literature on FDI (Egger and Pfaffermayr, 2004; Portes and Rey, 2005; Burchardi et al., 2019) which in essence constitutes a cross-border firm or subsidiary relocation. The model's intellectual baseline is that larger counties are expected to receive greater relocation flows due to gravitational force while two counties further apart have lower relocation flows. Moreover, firm-specific costs are captured in the gravity model's distance variable to overcome the fact that data on costs associated with firm mobility are not available.

Relocating firms in monopolistic competition incur considerable monetary moving costs and, possibly even more importantly, indirect costs such as costs to find new employees or establishing a new network and obtaining information. In contrast to the assumptions of a choice model, firms are assumed to have greater knowledge about geographically closer markets and locations. In the firm mobility literature, Conroy et al. (2016) and Pan et al. (2020) use a gravity-like model, albeit without considering the key determinant of the gravity model - the distance. However, Pan et al. (2020), suggest its usage for future studies. Hence, this Chapter is the first to analyse firms' relocation explicitly modelling the costs using distance as a proxy.

The theoretical justifications for the gravity model have been provided by Anderson (1979) and Anderson and Van Wincoop (2003) among others. In terms of the model's theoretical foundation, a profit maximizing individual firm k decides to relocate from location i (origin) to location j (destination) if the expected return R in the destination is greater than the expected return in the origin plus the costs of relocation C as a function of distance d:

$$E(R_j^k) > E(R_i^k) + C(d_{ij}) \tag{3.2}$$

When the above condition is satisfied and the firm relocates, the variable M_{ij}^k being equal to 1 (0 otherwise) is defined. By aggregating individual movements by county and employing a general gravity-type model specification, M_{ij}^k can be expressed as:

$$\Sigma M_{ij}^k = M_{ij} = f(\Delta l_{ij}, d_{ij}) \tag{3.3}$$

where i = 1, 2, ..., 401; j = 1,2, ..., 401 (with $i \neq j$) and l = vector of regional characteristics of the origin i and destination j and d_{ij} is the distance between i and j. I conduct the classical gravity specification expressed in log-linear form:

$$\ln(M_{ij,t}) = \alpha + \ln(l_{i,t}) + \ln(l_{j,t}) + N_{ij} + \ln(d_{ij}) + T_t + \gamma_i + \phi_j + \epsilon_{ij,t} \tag{3.4}$$

where $M_{ij,t}$ refers to gross relocation flows from county i to county j in year t. The dependent variable enters the model in log form to smoothen its distribution as well as to allow for the coefficients to be interpreted as semi-elasticities. In order to assess whether the same location factors are important for firm births as well as for relocating firms I use the same location characteristics l as presented in models 1.1 to 1.4 (population densities, service ratio, industrial ratio, gross income, price index (change), universities, research institutes). The specification for Model 2 (Equation 3.4) captures push and pull factors. Consequently, $\ln(l_{i,t})$ encompasses location characteristics for origin i (push factors), while $\ln(l_{j,t})$ are the corresponding location characteristics in destination j (pull factors) in year t.

N_{ij} is a dummy variable capturing contiguity equal to one if i and j are neighbours. The key variable to proxy costs of relocation d_{ij} is the linear distance in kilometer among the centroids of the counties. T is a year fixed effect that controls for common time shocks and general trends. γ_i and ϕ_j are time-invariant origin- and destination-specific fixed effects controlling for any unobservable county-specific factors affecting relocation and eliminating biases due to multilateral resistance (Bertoli and Moraga, 2013). $\epsilon_{ij,t}$ is the error term.

The second gravity model (Model (3), Equation 3.5) is a variant of Model (2) using a differential approach between origin and destination counties in order to investigate the differences in location characteristics (Conroy et al., 2016). This is a useful exercise as it can be expected that most firms make their decisions by evaluating their current location vis-à-vis other options with the new location reflecting the "winner" of these considerations and therefore capturing the firm's revealed locational preferences.
The empirical representation for the panel data can be expressed as

$$\ln(M_{ij,t}) = \alpha + \Sigma_{l=1}^{L} \Delta X_{l,ij,t} + N_{ij} + \ln(d_{ij}) + T_t + \gamma_i + \phi_j + \epsilon_{ij,t} \tag{3.5}$$

where $M_{ij,t}$ refers to the gross relocation flow from county i to county j measured as the log of the count of firms relocating from i into county j in year t. $\Delta X_{l,ij,t}$ is the difference in the location characteristics between origin county i and destination county j (destination minus origin) across the set of l empirical variables in year t. N_{ij}, $\ln(d_{ij})$, T_t, γ_i, ϕ_j, and $\epsilon_{ij,t}$ are the same as in Model (2).

This Chapter relies on both using OLS as well as Pseudo Poisson Maximum Likelihood (PPML) as estimation methods for the gravity model. Comparing results of these two empirical exercises allows interesting conclusions as to the roles of zero-observations. As highlighted by Silva and Tenreyro (2006), the most common practice in empirical applications of the gravity model has been to take natural logarithms and to estimate the model by OLS regressions. However, the literature, in particular on international trade, has developed several empirical solutions to deal with zero observations, that is non-flows. One possible and commonly used solution is the PPML model as proposed and discussed by Silva and Tenreyro (2006).

In the PPML, possible biases due to the amount of zero observations in the dependent variable are corrected while also accounting for heteroscedasticity. Further, Monte-Carlo simulations show that the estimator performs well in spite a large proportion of zeros (Yotov et al., 2016, p. 20) and the validity of the estimator does not depend on very strong assumptions of the distribution of the data as for example would be the case for a zero-inflated model (Santos Silva and Tenreyro, 2022). Overall, the PPML estimator is widely accepted to be very well suited for gravity estimations.

3.6 Empirical Results

3.6.1 Firm birth locations

Econometric results on German firm births between 2008 and 2017 are presented in Table 3.3, where Columns (1) and (2) correspond to the models based on firm birth per 1,000 inhabitants. Model 1.1 shows the results of the fixed-effects estimation and Model 1.2 displays the results of a pooled OLS (estimation without fixed effects). The results of the pooled OLS mainly serve as a benchmark to assess level effects against the fixed effects models that captures the sensitivity to changes in county characteristics and to assess which county characteristics are indeed absorbed by the fixed effects.[11] Column (3) and (4) refer to Model (1.3 fixed effects and 1.4 pooled OLS) with the total number of firm birth as dependent variable.

General location-specific factors

Population density is used to proxy for agglomeration effects (Rosenthal and Strange, 2008) and to partially control for market size and accessibility (Arauzo Carod, 2005). That is why a positive effect for population density on firm birth is expected (see Hypothesis 1 in Section 3.3.3). With regard to Model 1.2 and 1.4, there is a positive and significant effect for agglomeration effects on digital firm birth as in line with the literature as well as theory-based expectations laid out above. However, when including time- and county- fixed effects that control for general time trends as well as county-specific characteristics (Model 1.1 and 1.3), the effect of agglomeration benefits stays significant but the coefficient becomes negative. That is, an above-average growth in population density (which implies an above-average growth in population) has a negative effect on firm birth.

While this switch in the signs may seem counter-intuitive at first, it overall shows that as shown by the pooled model densely populated counties offer a host of agglomeration externalities conducive to firm birth. The fixed effects model instead reveals that the change in population density – as the effect is identified by difference in the growth of population density, does not have an effect on digital firm birth. Thus, above-average growth is probably driven by general labour market effects such as strong industry players that offer attractive jobs or various amenities of

[11] The pooled OLS model seems inappropriate to serve as a baseline model. The fixed effects model estimation appears superior because i. it seems unlikely that county characteristics are not randomly distributed and ii. the fixed effects are individually significant and increase the R^2. Results of a Hausman-Test show that fixed effects is chosen over random effects.

Table 3.3: Regression Results Model (1) Firm birth (OLS)

	Firm birth per 1,000 inhabitants (ln)		Total firmbirth (ln)	
	(1.1)	(1.2)	(1.3)	(1.4)
Population density (ln)	-0.192***	0.003***	-1.612***	0.083***
	(0.031)	(0.001)	(0.486)	(0.000)
Industrial ratio	-0.001	0.0005**	-0.0002	-0.002
	(0.001)	(0.0001)	(0.006)	(0.054)
Service ratio	-0.001***	0.0003***	-0.015***	-0.008***
	(0.0005)	(0.0001)	(0.005)	(0.000)
Gross income (ln)	0.005	-0.066***	-0.231	0.031
	(0.024)	(0.006)	(0.364)	(0.736)
Firms per 1,000 inhabitants (ln)	0.183***	0.157***	2.171***	0.613***
	(0.010)	(0.005)	(0.144)	(0.000)
Price index	0.002***	0.001***	0.030***	0.007***
	(0.0004)	(0.00004)	(0.005)	(0.000)
Price index change	-0.002***	-0.002***	-0.028***	-0.009***
	(0.0003)	(0.0001)	(0.005)	(0.000)
Metropole (G)	0.043	0.019***	2.977***	0.159***
	(0.038)	(0.005)	(0.384)	(0.00004)
Neighbour is metro (NG)	-0.583***	0.001	-4.598***	0.089***
	(0.088)	(0.001)	(1.354)	(0.00001)
Firm birth lag (ln)	-0.009***	0.0005	-0.142***	0.591***
	(0.001)	(0.001)	(0.019)	(0.000)
Universities	0.005**	0.0004	-0.043*	0.050***
	(0.002)	(0.0004)	(0.024)	(0.000)
Research Institutes	0.00002	0.002***	-0.031	0.007***
	(0.003)	(0.0003)	(0.028)	(0.005)
County FE	Yes	No	Yes	No
Time Fe	Yes	No	Yes	No
Observations	4,010	4,010	4,010	4,010
Adjusted R^2	0.875	0.767	0.889	0.798

Notes: significance levels are: ***p < 0.001; **p < 0.01; *p < 0.05; all standard errors clustered by county (in parentheses next to coefficients); dependent variables as logarithms plus 1.

dense cities. As reflected in the knowledge spillover theory of entrepreneurship (Audretsch et al., 2008; Acs et al., 2009), new innovative firms as they are part of the employed dataset thereby originate in the commercialisation of ideas in incumbent firms or universities. These new ideas need time to develop into marketable innovations, what makes it reasonable to expect a time-lag from a growth in population until a growth in firm birth.

Similarly, there are significantly more digital firm births in cities above 500,000 inhabitants than elsewhere (see Model 1.2 and 1.4). However, once controlling for unobservable characteristics of such metropolis (Model 1.1 and 1.3), there still is a positive effect in absolute terms, but not relative to the population. Therefore, an above average firm birth activity is not solely driven by the quantity of the population, but by unique, unobservable characteristics of metropolitan cities that oftentimes come with greater density like for example a cultural setting or entrepreneurial culture.

These unobservable advantages seem to spill over into neighbouring counties, as there is significantly more firm birth (Model 1.4) for neighbours of metropolis cities, but the effect vanishes with the inclusion of the fixed effect (Model 1.1 and 1.3). This is evidence that core cities' networks and knowledge are accessed at lower costs. Moreover, entrepreneurs might be too risk-averse to rent expensive offices in their first years of uncertain income streams.

Industry-specific factors

Localisation economies have a positive effect on firm birth. For each digital firm per 1,000 inhabitants, there is 0.18 digital firm births per 1,000 inhabitants, or in other words, five digital firms bring about one additional firm birth. There are two possible explanations for the observed dependency. First, co-location lowers the costs of starting a business and allows access to a more diverse range of inputs and complementary goods (Feldman et al., 2005).

With a large number of small-sized firms (I assume this is the case as laid out above), the chance of employees leaving a firm and starting their own business increases (Pijnenburg and Kholodilin, 2014). Therefore, a part of the higher firm birth rates with higher co-location potentially occurs from spin-offs. Further, the number of firm births when lagged by one year is significant and negative. This means that the average county does not register steadily positive growth rates in the digital sector, but high growth rates is a phenomenon in very few cities.

However, the models also reveal that a higher service ratio is associated with less digital firm birth in absolute and relative terms. This could mean that a very high regional specialisation in one service sector (for example Frankfurt in banking or Düsseldorf in consulting) does not offer enough industrial diversity for digital firm birth. Considering the fact that the service ratio includes ICT services as well, this might be indicative evidence that digital firms seek proximity to a broad number of recipients which are not other service sector firms. This is in contrast to the fact that there is no such effect for the industrial ratio. Second, with increasing co-location of similar firms, a large labour pool of specialized workers - crucial for digital businesses - is available. This agglomeration advantage in addition to the local amenities exceed the disadvantages of higher prices (Rossi and Dej, 2020). This is also what I find in my model: In counties with 5 % higher prices for apartment rents than German average there is one firm birth per 1,000 inhabitants, while a 5 % growth in prices (relative to the German average) leads to one less firm birth per 1,000 inhabitants. That is, higher price levels are accepted by entrepreneurs while strong price growth has a negative impact on firm birth. However, gross income as a measure of overall wealth does not have a significant effect. In sum, this is indicative that co-location benefits outweigh the costs.

Digital firm birth activity was expected to be high in cities where firms derive advantages from agglomeration benefits, that is firms show a strong preference towards cities (Hypothesis 1). Overall, results show a high level of agglomeration benefits is conducive to firm birth, and that digital firm birth activity is mostly driven by strong co-location of similar firms and the regional knowledge base.

Another important finding is that universities contribute significantly to higher firm birth rates providing support for Hypothesis 2 (Section 3.3.3). For each additional university there is 0.5 firm births per 1,000 inhabitants. For example, an average county with a university has 38 digital firm births on average, while a county without a university has only 11. One link between universities and start-up activity is that innovative students or college employees found spin-offs because the expected value of commercialising an idea is higher for the individual than the expected income offered through employment in an established firm (Pijnenburg and Kholodilin, 2014).

Another link is that university-bound knowledge may breed product development particularly by young digital firms while regional knowledge is crucial for start-up rates due to spillovers' spatial limits (Bade and Nerlinger, 2000). The impact of research institutes is significant in the pooled model, while there is no significant effect in the fixed effects model. This shows that a general presence of research institutes is conducive to firm birth, but an increase does not imply more firm birth in the same year. Hypothesis 2 is therefore partly confirmed. This indicates that a mere geographical presence of knowledge-producing institutions is not sufficient for digital firm birth, but rather that transfer channels and potential labour market effects differ between research institutions and universities.

3.6.2 Relocation

Table 3.4 and 3.5 show the results for the gravity models laid out in Equation (3.4) and (3.5). In Model (2), absolute origin and destination characteristics are included. A positive/negative coefficient in the origin/destination characteristic implies a higher/lower flow of out-migrating/in-migrating firms. In Model (3), the variables display the difference between origin and destination. Thus, a positive/negative coefficient implies higher/lower values in the destination than in the origin.

Similar to the results on firm birth, the models are presented with fixed effects (Column 2.1) and without fixed effects (Column 2.2), respectively. In the fixed effects models which controls for location-specific characteristics and common time trends which affect all locations none of the coefficients for regional characteristics have significant explanatory power in comparison to the models without fixed effects. This result in its own right sheds light on the lack of relevance of regional characteristics up and beyond their idiosyncrasies for firms' relocation decisions.

That is, why the fixed effects model employed here generates considerably less significant effects on the vector of regional characteristics explicitly controlled for in the model in contrast to e.g. Kronenberg (2013) who employs a pooled OLS estimation without fixed effects. Nonetheless, the model's remaining significant determinants - strikingly mostly industrial factors - can be interpreted as relevant factors for inter-regional relocation when controlling for county characteristics. As there are many unobservable regional characteristics I prefer to focus on the fixed effects model for the interpretation of the results.

Industry-specific factors

Results show that other than expected, except for the numer of digital firms per 1,000 inhabitants (in the destination) and universities (in the origin), none of the regional characteristics have a statistically significant effect on relocation. Therefore, the industrial- and knowledge base is most important. An urban 'upgrading' in e.g. amenities is of limited use to foster firm birth, as soft factors are not identified as pull factors.

Both gravity models show a significant 'gravitational' effect for localisation as measured by

Table 3.4: Regression Results Model (2) Gravity Model (PPML)

	M_{ijt} (2.1)		M_{ijt} (2.2)	
Population density (ln) Origin	1.19772	(1.05330)	0.12288***	(0.01350)
Population density (ln) Destination	0.26170	(0.99375)	0.12734***	(0.01331)
Industrial ratio Origin	0.02284	(0.01556)	-0.03132***	(0.00166)
Industrial ratio Destination	0.01307	(0.01609)	-0.02101***	(0.00165)
Service ratio Origin	0.02151	(0.01238)	-0.00050	(0.00086)
Service ratio Destination	0.00223	(0.01210)	-0.00702***	(0.00092)
Gross income (ln) Origin	-1.48138	(0.83224)	2.04666***	(0.11828)
Gross income (ln) Destination	0.68121	(0.81643)	1.25158***	(0.11832)
Price index Origin	-0.00473	(0.01103)	0.00997***	(0.00060)
Price index Destination	0.01612	(0.01055)	0.00563***	(0.00063)
Price index change Origin	0.00218	(0.00962)	-0.00952***	(0.00102)
Price index change Destination	-0.01577	(0.00942)	-0.00726***	(0.00108)
Firm birth (ln) Origin	-0.07896	(0.04685)	0.16464***	(0.02044)
Firm birth (ln) Destination	0.02792	(0.04455)	0.33232***	(0.02124)
Firms per 1,000inhabitants Origin	-0.13080	(0.09183)	0.00020	(0.00011)
Firms per 1,000inhabitants Destination	0.20865 *	(0.08660)	0.00014	(0.00012)
Research Institutes Origin	-0.00839	(0.04625)	0.00282	(0.00249)
Research Institutes Destination	-0.05027	(0.04794)	-0.01771***	(0.00270)
Universities Origin	0.07386 *	(0.03258)	0.10703***	(0.00360)
Universities Destination	0.01235	(0.03536)	0.13669***	(0.00385)
Neighbour county	1.21148 ***	(0.04223)	1.24792***	(0.03039)
Distance (ln)	-1.15491 ***	(0.01647)	-1.03266***	(0.01164)
Time Fixed Effect	Yes		No	
County Fixed Effects	Yes		No	
Num. Obs.	1432818		1443600	

Notes: significance levels at ***p < 0.001; **p < 0.01; *p < 0.05, standard errors in parentheses; dependent variable as log-link

Table 3.5: Regression Results Model (3) Gravity Model Difference Approach (PPML)

	M_{ijt} (3.1)		M_{ijt} (3.2)	
Population density	0.00013	(0.00029)	-0.00015***	(0.00001)
Industrial ratio	-0.00130	(0.01076)	0.00215*	(0.00100)
Service ratio	-0.00733	(0.00791)	-0.00117	(0.00061)
Gross income	0.00025	(0.00022)	-0.00003	(0.00003)
Price index	0.01662	(0.00900)	-0.00651***	(0.00065)
Price index change	-0.01464	(0.00808)	0.00465***	(0.00110)
Firm birth	-0.00021	(0.00034)	-0.00064***	(0.00015)
Firms per 1,000 inhabitants	0.17086 *	(0.06783)	0.09770***	(0.02743)
Research Institutes	-0.02315	(0.03586)	-0.02936***	(0.00286)
Universities	-0.01905	(0.02846)	0.05375***	(0.00467)
Neighbour County	1.21144 ***	(0.04226)	1.62317***	(0.02775)
Distance (ln)	-1.15495 ***	(0.01648)	-1.03624***	(0.00980)
Time fixed effects	Yes		No	
County fixed effects	Yes		No	
Num. Obs.	1432818		1443600	

Notes: significant levels at ***p $<$ 0.001; **p $<$ 0.01; *p $<$ 0.05, year, origin and destination fixed effects included; standard errors in parentheses; variables coded as differences between origin and destination; firm birth and firms per 1,000 inhabitants refer to digital firms; dependent variable as log-link

firm density (digital firms per 1,000 inhabitants) and thus same-sector density can be identified as a relevant pull factor. Model (2) predicts that with one additional digital firm per 1,000 inhabitants relocation inflows increase by 23 percentage points. Thus, Model (3) partly provides support for Hypothesis 3, that is, firms do not necessarily move.

This result is different from Stam (2007) and Nguyen et al. (2013) investigation of manufacturing firms. In contrast to manufacturing, digital firms are more likely to cluster and seek out competition. Furthermore, there is no significant effect for salaries or housing prices. In line with Kronenberg (2013) and Rossi and Dej (2020), this finding implies that digital firms do not necessarily adopt pure cost minimisation, but choose counties where they can benefit from the availability of high-skilled workers. This is comparable to the above finding for digital start-ups, indicating an industry-specific behavior which is independent of firm maturity.

Moreover, my results give weak indication that digital firms move from diverse into more specialized cities seeking proximity to their competitors. For manufacturing, Duranton and Puga (2001) find firms to innovate in diversified cities and then switching to mass production in localized cities. Although digital firms are usually not producing physical goods and economies of scale are limited in many digital branches, it seems reasonable that digital firms do reach some point of process standardisation and stable growth path for which locational diversity is less of a requirement. That is, knowledge bound to diverse cities and their universities can become less relevant for digital firms later in their lifecycle. For every university, the outflow of firms increases by 107 % (see Model (2)). When assuming that universities contribute to a high and diverse output of firm births in different sectors and thereby to a diverse local economy, this finding provides further support for the hypothesis that firms move from diverse into more specialized places.

An alternative explanation may be that firms seek direct competition with others as innovation and knowledge in software, for example, are hard to be legally protected via patents thus rendering spillovers, networking and shared work and customer flows easier and more attractive. Consequently, while firm birth seeks knowledge bound in universities, relocates seek industry-specific knowledge and competition.

In sum, digital firms show a strong tendency for clustering and co-location with other digital firms. This is a self-reinforcing mechanism that contributes to higher firm birth and inflows of relocating businesses. Therefore, the expectation that agglomerated areas receive higher inflows (Hypothesis 3) cannot necessarily be confirmed since benefits for firms are rather driven by knowledge and industry rather than pure population density.

Spatial mobility pattern

Model (2) and (3) reveal a very strong regional persistence of digital firms. This supports the indicative findings from the descriptive statistics. In light of the high number of zero-flows, this result is even more striking. Expecting transaction costs to increase with distance, counties being in closer geographical proximity show higher relocation flows than counties further apart. The results show that with a one unit increase in the $\ln(distance)$ between county i and j the average sizes of the relocation flow from i to j decreases by the factor of 0.31. Compared to all other relocation flows, flows between neighbouring counties are expected to be 236 % higher. In other words, counties receive more than twice as many relocates from their neighbouring counties than from others. This confirms the expectation from Hypothesis 4 that relocation flows decrease with distance. As shown in Section 3.4.3, the flow between the City of Munich and Munich County account for the biggest relocation flows. In light of the dominance of these observations in the sample, I repeat the empirical exercise above excluding this pair. My results still hold (see Appendix Table A1.4 and Table A1.5) indicating that regional persistence and industrial path dependence are crucial for all German counties and results are not solely driven by the "Greater Munich effect".

Moreover, the results are consistent with Conroy et al. (2016) who also find a strong neighbour effect for interstate relocation for US manufacturing firms. Furthermore, the results are in line with Knoben (2011) who finds that young ICT firms, which are highly dependent on outside resources remain within the local economy since costs do play a significant, prohibitive role in relocation choices.

3.7 Robustness

Several robustness test have been conducted. To assess possible multicollinearity issues in the specification, a correlation table is presented in Table A1.2. Due to the high correlation of universities and research institutes (0.8), a variance inflation factor (see Table A1.3) has been calculated. Results show that none of the values is above ten, which indicates that multicollinearity is not a concern for the regression results (Wooldridge, 2015). Moreover, when taking out either of these variables from the model or when aggregating the two variables the estimations yield the same results and R remains the same. Next to confirming robustness, these results indicate that the findings presented above indeed warrant a differentiation between the different types of knowledge these institutions offer for digital industries.

For the robustness of the gravity models, the key concerns certainly are the large number of zero-flows, as in fact only 0.7 % amount of observations in my sample are positive. When including these zero observations in the estimation as done above conceptually, I am estimating county-pairs with firm mobility relative to non-integrated counties that is those ones that have no firm mobility between them (the majority of observations). The advantage of this estimation is that it allows to understand factors which explain why certain counties have firm mobility at all and the determinants which "switch on" firm mobility. This is similar to the extensive margin in the international trade literature (Helpman et al., 2008; Chaney, 2008) and the literature on FDI (Blum et al., 2020), and thus my preferred specification.

To analyse the relevance of zero-observations for my findings, I repeat all gravity estimations as a Negative Binomial (NB) as an alternative to the PPML and find consistent results. Furthermore, I repeat the empirical exercise of the main specification this time excluding all county-pairs without any flows, thereby effectively estimating the intensity of firm mobility between counties with any firm migration. I conduct this exercise using the three standard methods in the literature PPML, NB and OLS (see Appendix A1.4 and A1.5). Results can be interpreted as a difference between the extensive and intensive margin in relocation flows for regional characteristics. Results of a specification excluding zero-observations and the one carried out above including all observations allows for re-pivoting the analysis towards focusing on the determinants within "mobility-relevant" counties rather than benchmarking "mobility-relevant" counties against those without any mobility at all.

Across all models, the distance between counties and contiguity are the predominant explanatory factors for the size of relocation flows between counties. Both variables are significant at a 99 % level and similar in magnitude to the specification analysed above. This is indicative of the fact that there are structural differences between receiving/sending counties and those which never have any mobility of digital firms which explain mobility which is in line with the persistence seen in the summary statistics. Seeing that both conceptual as well as methodological approaches render consistent results yields further support for the analysis presented above.

As elaborated in Section 3.5, the inclusion of origin- and destination fixed effects in the panel model above controls for multilateral resistance, as commonly done in the international trade literature using gravity models. Moreover, origin-time and destination-time fixed effects to fully account for multilateral resistance are commonly employed (Olivero and Yotov, 2012; Head and Mayer, 2014, p.151).

In the relocation context, these fixed effects can be seen as a barrier to move that a firm faces with all its possible new locations. In other words, the estimation captures relative changes in counties attractiveness while controlling for all observable and unobservable country-specific variables that vary in the respective dimension (Yotov et al., 2016, p.19). In this case, the inclusion of origin-time and destination-time fixed effects leaves variation in contiguity and distance. The results in Table A1.6 show that the size and significance of the coefficients remain consistent with the baseline estimations in Table 3.4.

Further, binary choice models capturing the probability to move have been estimated (Appendix Table A1.7). The advantage of such models over continuous models as used in this Chapter is that individual firm characteristics are controlled for. That is one of the reasons why they have been used in the literature so extensively (Kronenberg, 2013; Nguyen et al., 2013). Results also identify co-location of similar firms as a significant pull factor. I do find that the spatial components of distance and flow into neighbouring counties are most significant across all models. Thus, there is reasonable ground to believe that the gravity model with its theory-based spatial dynamic reflects and encapsulates the results obtained by a choice model. Moreover, this robustness check confirms that my findings in spite of a different empirical strategy are comparable to Kronenberg (2013) and Nguyen et al. (2013).

In order to test the feasibility of the data selection of digital firms, the relocation analysis has been conducted with a reduced sample by only using the firms that have been selected as ICT via the NACE code (Table A1.8 and A1.9). This sample is more restrictive and leaves 3,600 origin-destination pairs with positive relocation-flows (instead of 10,108 in the full sample). This selection absorbs some heterogeneity of the firm landscape: I assume that firms are more similar in their business models referring back to the example of 31.8 % of new businesses being registered in ICT but 66 % state to operate on a digital business model (Kollmann et al., 2020). Results show, contrary to the full sample selection, that higher firm birth rates in counties results in higher outflows of ICT firms. This shows that, the more homogeneous the market, that is higher competition as less diversity, the more firms leave. Therefore, a diverse firm landscape of the digital economy can be advantageous for a county to prevent firms from leaving.

Additionally, a complementary analysis has been performed on NUTS 2 aggregation (see Tables A1.10 and A1.11). This aggregate includes firms that relocate over longer distances and therefore tolerate higher relocation costs (in comparison to NUTS 3) and increases the share of positive flows to 29 %. The striking difference between the two geographical aggregations is that the gross income in the origin is significant and negative across all tested specifications and the number of universities in the destination is significant and positive.

Drawing on the difference to the results from the NUTS 3 aggregation that shows a clearer picture on the role of universities and knowledge diffusion: As NUTS 3-units show a higher outflow of digital firms with more universities, NUTS 2- units show a higher in-flow of digital firms. This may indicate that access to knowledge is a striking advantage that firms accept higher costs for, especially in very early stages. Further, firms accepting higher costs of movement relocate into regions with lower gross income that is they follow a price mechanism here.

In sum, the key findings of this Chapter are robust to various specifications. One caveat is that the model may suffer from endogeneity issues, such as selection into locations based on characteristics. The best to do to remedy is to employ a range of fixed effects to identify the effect. Moreover, it is reasonable to assume that my results indicate a lower bound for the “real” effect should selection be an issue therefore not posing a substantial threat to the empirical strategy of the Chapter.

3.8 Conclusions

This study investigates the determinants of firm birth location and relocation choices of young digital firms in Germany from 2008 to 2017. The main objective of the chapter is to shed light on the importance of regional characteristics for firm birth and relocation of firms, by asking if both types of location choices are driven by the same determinants. Results reveal that the employed set of regional characteristics is conducive to firm birth, while only very few of these regional characteristics are significant when seeking to explain firm relocation. For the latter, the major explanatory factor is distance, that is the cost of relocation as well as industry-relevant factors, such as density of digital firms. Therefore, I conclude that agglomeration effects do in fact matter for digital firms, as firms tend to stay in the region where benefits arising from agglomeration can potentially be accessed. For digital firm birth in Germany, results are in line with the theoretical prediction of Hypothesis 1 with firms displaying a preference towards cities (Van Oort and Atzema, 2004; Trippl et al., 2009). By employing pooled as well as fixed effects models, the results indicate that high levels of agglomeration benefits are conducive to firm birth in general. Nevertheless, an above average growth that increases the benefits is not conducive to new firms in a given year. However, results also show advantages for digital firm birth arise mostly from industry-specific benefits such as other digital firms. Therefore policy makers should focus on enhancing industry-specific factors. Additionally, local knowledge available through universities and research institutes is conducive to new firm formations (Hypothesis 2), although research institutes play a minor role.

In terms of relocation, Hypothesis 3 is that counties with higher agglomeration benefits such as a specialized high-tech labour market and the potential for Information Technology (IT)-specific knowledge spillovers attract more relocations. The analysis shows that relocation inflows are higher in cities with high digital firm densities. The results suggest that firms expect that negative effects of co-location such as congestion, competition and higher prices do not outweigh the benefits and advantages derived from co-location of similar firms such as access to industry-specific knowledge and a specialized labour pool.

Further, the forth hypothesis was that most relocations occur between geographically proximate or contiguous counties where moving costs are low while access to locally bound factors such as local customers, suppliers or networks remains relatively low-cost. Previous research indicates that young digital firms are highly mobile as they are less exposed to incur sunk costs in the relocation process (Sleutjes and Beckers, 2013; Esteve-Pérez et al., 2018). With the usage of the gravity model, this study provides evidence that relocation costs play a significant role for digital firms. Thus, I conclude that digital firms are not as footloose even in their youth as one might expect. The distance between firm birth location and relocation destination is the predominant explanatory factor for relocation when controlling for several regional characteristics in the employed fixed effects model. Therefore, digital firms predominantly stay in the region where they were born initially or in the direct vicinity. This is why regional economic patterns remain very persistent over long time spans.

The results of the study are particularly relevant for policy makers trying to foster local economic growth by attracting digital firms. The study reveals that strong co-location of similar firms deepens regional specialisation, since industry concentration is very conducive to attracting digital firms as well as relocating businesses. The study also shows that digital entrepreneurship is a regional phenomenon when considering the medium term. That is, cities offer a breeding ground for new digital businesses and in the medium run, the surrounding counties benefit once these firms decide to leave their birth towns. Moreover, these results imply that firms are willing to access networks and knowledge that is bound in the region after relocation. Thus, competitiveness-improving effects are not limited to a certain region but spread across NUTS 3 borders to a limited extend (Pijnenburg and Kholodilin, 2014). Once a location decision has been made, it is very likely that the firm will be staying in its region of origin which leads to reinforcing spatial patterns of digital firms. Political measures that target start-up rates can have a positive, long term effect on neighbouring counties.

Chapter 4

Shoulders and Shadows of Giants – Intra-regional Distribution of the Digital Industry in Germany

Abstract

This Chapter investigates digital firm birth activity in municipalities in the urban hinterland of core cities in Germany. I conduct panel fixed effect regressions for monocentric and polycentric urban labour market regions covering the years 1995 to 2017. The digital industry's regional distribution is significantly shaped by the closest core cities: municipalities in monocentric urban regions profit from urban population growth and universities' general knowledge. Municipalities in polycentric urban regions, however, are affected by industry-specific externalities, that is, an above-average growth in the share of firm birth of their closest urban cores. Overall, agglomeration externalities experience spatial decay relative to the core size with all regions benefiting from their own industry-enhancing urbanisation externalities as captured by population growth and universities.

4.1 Introduction

Entrepreneurial activity as pronounced in the birth of new, innovative firms, is a predominantly urban phenomenon where firms derive agglomeration benefits. Because administrative borders do not cut off agglomeration externalities and spillovers, high levels of entrepreneurship can be advantageous for whole regions, resulting in regional persistence of firm birth patterns (Pijnenburg and Kholodilin, 2014; Fossen and Martin, 2018; Stuetzer et al., 2021). Nevertheless, it remains unclear whether benefits deriving from a city as the nucleus of regional development disperse homogenously in space.

This chapter investigates digital firm birth patterns from 1995 to 2017 in the surrounding municipalities (LAU2 Regions, Gemeindeverbände) of German core cities.[1] It investigates firm birth in small communities serving their bigger neighbours' labour markets, and in return receiving income flows (Parr, 2014). Thereby, the analysis is not limited to high rank first-cities, but includes second- and third tier cities as their impact on economic growth has been given attention in the literature (Dijkstra et al., 2013; Camagni et al., 2015). The analysis of the digital industry on the municipality level complements the analysis of NUTS 3 regions, as predominantly used to assess regional development (Pijnenburg and Kholodilin, 2014; Fossen and Martin, 2018). The digital sector is particularly well suited to proxy entrepreneurship more generally as it has an inherent tendency towards geographical agglomeration (Moretti, 2012) and is a strong start-up sector. Further, the relatively new sector is characterised by strong growth rates and complements many other sectors. In addition, due to the knowledge-intensive nature of digital products and services, there are little natural advantages and relatively low sunk costs. At the same time, the sector depends on agglomeration externalities and thick labour markets in particular. Thus, firms' location choices are very sensitive to externalities and largely dependent on the existing regional knowledge base. Due to these characteristics, the digital sector has been the target for many policy interventions seeking to foster regional growth.

This chapter contributes to a deeper understanding of the spatial dimensions of agglomeration externalities, competition, and dispersion effects within urban regions. I focus on answering the following three main questions: First, what are relevant location factors for smaller municipalities within core city regions that contribute to the capitalisation of entrepreneurship capital? Second, I investigate the relation of the municipalities' firm birth dynamics vis-à-vis

[1] Core cities are cities with more than 100,000 inhabitants and a surplus of inbound commuters. Additionally, the main commuter flow does not come from the neighboring center, as defined by BBSR (2022b). The terms core cites and centers are used interchangeably.

the dynamics within the respective core cities. This is especially relevant for policy makers in small and medium-sized cities that function as regional development engines outside metropolitan regions that have an outstanding significance for a larger surrounding area, according to the Federal Ministry of Housing, Urban Development and Construction (Federal Ministry of Housing and Construction, 2022). Third, I investigate the industry's regional spatial distribution in the long run by estimating whether the distribution of entrepreneurship capital differs between Monocentric Urban Region (MUR) and Polycentric Urban Region (PUR) (whose labour market regions host more than one core city) due to presumably overlapping agglomeration effects of two core cities.

The political importance of supporting PURs is emphasized in key strategic documents of the European Union (EU) (European Union, 2011), namely the European Spatial Development Perspective (European Commission, 1999), the EU Territorial Agenda 2020 (European Commission, 2011) and the "Pact of Amsterdam" which established an EU Urban Agenda (ESPON, 2017). The morphology of German city-regions offers unique patterns for analyzing municipalities given their location in the vicinity of core cities: Cities are well connected due to short distances compared to the US and economic activity is less centralized than in other European countries, most notably France.

For the empirical analysis, I use panel fixed effects regressions for firm birth in German municipalities belonging to a core cities' labour market region from 1995 to 2017. The share of firm birth in a municipality relative to all firms born in the respective labour market region serves as the dependent variable. Hence, the results reflect relative advantages of small municipalities surrounding cities. Results show that municipalities' individual characteristics - population growth and hosting universities - have a positive effect on increasing the share of digital firm birth start-up activity by 0.38 % and 0.95 % respectively. In line with the literature, the share of firm birth decays by approximately 0.01 % with each additional kilometer of distance to the urban core. Further, firm birth activity is shaped by the morphology of the region: Municipalities in MURs are exposed to strong competition against their cores in terms of industry-enhancing factors. In contrast, municipalities in PURs gain advantages from a growing industry in the center closest to them. The chapter will proceed as follows. Section 4.2 gives an overview of the related literature on agglomeration effects, entrepreneurship capital and the literature on MURs and PURs. Further, it presents the aim and hypotheses of the Chapter. Section 4.3 presents the data and descriptive statistics. Section 4.4 describes the empirical strategy, while Section 4.5 presents the results. Section 4.6 concludes and provides policy implications.

4.2 Related Literature and Hypothesis

4.2.1 Agglomeration and Entrepreneurship Capital

Local start-up rates in knowledge-intensive industries like ICT are higher in larger cities and surrounding areas (Bade and Nerlinger, 2000; Audretsch et al., 2012; Van Oort and Atzema, 2004; Pijnenburg and Kholodilin, 2014). This is mainly due to advantages deriving from agglomeration externalities such as sharing infrastructure, matching effects of thick labour markets and learning through knowledge spillovers (Duranton and Puga, 2001; Puga, 2010)

The learning channel is particularly important for digital companies as new knowledge created in both public and private knowledge institutions as well as industry competitors manifests itself by additional firm birth (Audretsch et al., 2008). As stated in the "Knowledge Spillover Theory of Entrepreneurship", an individual will start a new business if the expected value of a piece of knowledge is higher for the individual than for a decision maker within an incumbent firm or university (Audretsch and Keilbach, 2007; Acs et al., 2009).

Nevertheless, local and regional start-up rates have been shown to be very persistent over time despite major economic disruptions (Fritsch and Wyrwich, 2017; Stuetzer et al., 2021). The explanation of this phenomenon lies in the entrepreneurship capital, a deeply rooted social acceptance to encourage and support startup activities through norms and values, strong formal and informal networks as well as high endowments of individuals willing to start a business (Audretsch et al., 2008). This is typically measured in firm birth rates. As the local culture drives entrepreneurial capital, it is tied to a region and locally bound (Audretsch and Keilbach, 2007; Fritsch, 2017).

Further, Fossen and Martin (2018) not only find entrepreneurship capital manifesting itself in regional startup rates, but also a spatial dependence to neighbouring regions in Germany. For the high-tech industry, a larger population nearby implies significantly higher startup rates in the short term. This is best explained by employees leaving successful employers to start an own businesses near home where they have good knowledge of local networks (Klepper, 2002). Therefore, big cities within a short distance facilitate the exchange of industry relevant knowledge and ideas as found in Fritsch and Wyrwich (2017); Fritsch and Aamoucke (2013, 2017), as well as Pijnenburg and Kholodilin (2014). The latter find spatial entrepreneurship spillovers for German NUTS 3 to extend over a range of about 50 km from the focus region. The reasons for these spatial interactions are twofold. On the one hand, increased startup rates enhance opportunities and lower the costs of starting a business, for example by providing access to suppliers and customers (Delgado et al., 2010). These advantages are accessible for neighbours, especially on NUTS 3 levels. One the other hand, high startup rates also increase the competition in neighbouring regions (e.g. capital investment), which would be a competitiveness-improving effect of entrepreneurship capital (Delgado et al., 2010; Pijnenburg and Kholodilin, 2014).

However, the above-mentioned authors mostly focus on regional growth mechanisms while paying little attention to the question of how these economic developments shape the studied regions in the long term. Theoretically, a region experiencing economic growth predominantly driven by externalities of a neighbouring region, should become large enough in the long run to produce ‘its own’ agglomeration externalities, at least some scale-externalities. Thus, we would either observe convergence resulting in a real expansion of externalities or even reach a point where competition between the players becomes very strong resulting in a dispersion of entrepreneurship capital (as discussed by Delgado et al. (2010) and Delgado et al. (2014)).

4.2.2 The morphology of the region: MURs vs. PURs

Scholars studying agglomeration effects in a core-periphery dynamic usually focus on intra-regional population distribution, that is, the sharing channel of agglomeration theory (Meijers et al., 2016; Volgmann and Rusche, 2020; Krehl and Siedentop, 2019). The difference to the literature presented above is cities and places are assumed to interact, where ‘performance’, in terms of population or industry growth, is dependent on the position within the region.

The literature on urban systems reflects a discussion similar to competition versus dispersion by the concepts of ’borrowing size’ and ’agglomeration shadows’. ’Borrowing size’ postulates smaller cities in larger urban areas inhabit more metropolitan functions (high-order economic, political and cultural features) than similar cities in less agglomerated areas. That is, a place borrows size when holding more metropolitan functions than its own size could normally support (Volgmann and Rusche, 2020). For example, Phelps et al. (2001) show how small cities around London can ‘borrow size’ by avoiding costs of agglomeration but still access specialized labour and the informal external economies.

Borrowing size is somewhat akin to the concept of spillover effects. In contrast, ’agglomeration shadows’ predict limited growth near high-tier cities due to higher competition crowding out economic activity (Meijers et al., 2016). Interestingly for this chapter, Volgmann and Rusche

(2020) find both borrowing size and agglomeration shadows for population distributions of German city-regions, showing that both effects co-exist in distinct regions.

Moreover, Volgmann and Münter (2022) compare differences in metropolitan growth of MURs with one dominant core with PURs having less spatial structural hierarchies. They argue that individual centers in PURs develop less agglomeration externalities due to lower concentration of population, cultural and political functions, but offer better quality of life due to lower negative externalities and congestion costs (Volgmann and Münter, 2022). In essence, it is assumed that agglomeration effects are distributed differently across MURs and PURs.

Empirical applications of these concepts focus on population distribution next to metropolitan functions rather than industry developments and labour market distributions (Meijers et al., 2016; Volgmann and Rusche, 2020; Volgmann and Münter, 2022). Ouwehand et al. (2022) investigate the impact of spatial structures on total factor productivity and conclude that – for European regions with similar urban populations – the urbanisation externalities derived from multiple city cores do not substitute for those achieved with a structure relying on singular, larger cities. This chapter contributes to the literature by putting the intellectual baseline of 'borrowing size' and 'agglomeration shadows' to an industry-context of entrepreneurship as the exposure to agglomeration externalities is a crucial input for digital, knowledge-intensive firms.

4.2.3 Aim and Hypothesis

The aim of the chapter is to shed light on intra-regional distribution of entrepreneurship capital apart from core cities by identifying characteristics of municipalities being attractive to firm birth next to urban cores.

First, I consider the interplay of entrepreneurship within regions on the municipality level (LAU regions). Entrepreneurship is measured by firm birth, because new firms are not constrained by previous location decisions and sunk costs. Therefore, they provide better information on the role and magnitude of agglomeration effects than existing ones (Gómez-Antonio and Sweeney, 2021). To shed light on the intra-regional distribution, I use the share of firm birth (firm birth in municipalities divided by the total firm birth of the region).

Second, Germany offers a dense system of cities, where, for example in the Ruhr-Area, several core cities are located within one large labour market region. In this Chapter, this is referred to as a PUR. Here, agglomeration effects and spillovers could be overlapping and causing different outcomes of competition and divergence compared to a MUR like e.g. Munich. Accordingly, the second aim of the Chapter is to identify differences in the distribution of entrepreneurship between mono-and polycentric dynamics. Thereby, I also tackle questions on the spatial range of agglomeration externalities.

The definition of the region is a fundamental determinant of the analysis' outcome. The smallest spatial scale applied in studies on regional interactions of entrepreneurship is NUTS 3 (county level). For the German case in particular, this mostly benchmarks core cities against their (urban) periphery (kreisfreie Städte vs. Kreise). This broad unit of analysis comes with certain disadvantages. Counties not being core cities can cover large areas and be heterogeneous within themselves, oftentimes hosting smaller, second-tier cities. Thus, the usage of LAU regions, as suggested by Volgmann and Münter (2022) offers more in-depth results for urban regions. Moreover, the principle of subsidiarity in Germany allows self-government of certain political and planning principles. Therefore, it is worthwhile investigating what relative advantages municipalities within the same labour market, i.e. similar locations within the system of cities, expose.

In general, I assume three major contributing factors: First, individual characteristics create small-scale agglomeration advantages for firms such as a sufficient population size (which proxies sharing) and a sufficient knowledge stock stemming from universities (a proxy for learning). Second, due to regional embeddedness, I assume the size of the next urban core and its knowledge stock to be a significant determinant of the spatial distribution. Third, as spillovers decay over distance (Rosenthal and Strange, 2003; Rice et al., 2006; Smit and De Groot, 2013), I expect advantages for municipalities located close to the core. Taken together, the first research question (Research Question 1) is: How do local factors and core city characteristics influence the intra-regional distribution of entrepreneurship capital?

Based on the empirical findings in the literature on 'agglomeration shadows' and 'borrowing size', it is likely that the impact of the core cities for the regional firm distribution differs between MURs and PURs. Therefore, Research Question 2 is: Does the intra-regional distribution of entrepreneurship capital differ in mono- and polycentric urban regions? For MURs, I expect a stronger spatial concentration of the industry than for PURs. In turn, PURs are expected to be more dispersed with agglomeration externalities which probably are greater in scope.

4.3 Data

4.3.1 Geo-coded firm level panel dataset

The advantage of the firm-data lies in the precision of individual firm birth locations over a long period. The analysis covers companies that entered the market between 1995 and 2017. The data provided by North Data (2019) originate from statutory publications of German corporations.

As there is no agreed-upon definition of the digital economy, a digital firm is defined as one that is information-technology driven and internet-based. I selected firms using NACE codes (similar to Weber et al. (2018)) covering general programming activities, software development, web portals, data processing, and the development of web pages, processing, hosting and related activities and web portals.[2] Yet, standard industry classification systems have limitations, especially industries which cross over traditional product categories, as is the case for digital firms (Oakey et al., 2001). Since digital business models complement many other sectors, firms may be registered in other NACE codes despite running a digital business model.

To identify these firms, the description of the company's main business area is used.[3] With the help of a word-search selection, firms not registered in the ICT sector but operating on a digital business model were added to the dataset. The resulting sample contains firms characterised by the core knowledge on which their competitiveness ultimately draws (Martin and Moodysson, 2013). The resulting dataset comprises 144,230 individual firms born between 1995 and 2017 and is aggregated to municipality level for the analysis.

4.3.2 Spatial units of analysis and location characteristics

The region is defined by the labour market (following the BBSR) because knowledge spillovers have shorter spatial ranges than labour market effects (Kerr and Kominers, 2015). The labour market defines the outer bounds of agglomeration externalities. Accordingly, the outer bounds of the labour market are where 25% to 50% of out-commuters commute to a center/supplementary

[2]NACE codes 62.01.0, 62.01.1, 62.01.9, 62.02.0, 62.03.0, 62.09.0, 63.11.0, 63.12.0.

[3]First, the description of the identified ICT firms has been analyzed and the most frequently used words related to IT and software has been identified. Then, these key-words are used to obtain those firms operating on digital business models with the help of several word combinations. Further, firms that only distribute their products via a webpage have been excluded. For those firms, key words related to "software development" needed to be included. As an example a firm that is registered in „ Placement of workers“ has been included in the sample, because the objective of the company is "the operation of a social networking platform for skills enhancement and marketing as well as the provision, brokerage and distribution of products and Internet-based services." A full list of keywords and in/exclusion of firms can be obtained from the author by request. Note that Keywords are in German.

area (BBSR, 2022b). A PUR is a (labour market) region that hosts more than one center, resulting in seven PURs.[4] A MUR hosts one center, which are 42 regions in Germany.[5] Within the regions, the unit of analysis is the municipality (see Figure 4.1). Since the goal of this Chapter is to assess the regional distribution of entrepreneurship apart from the core cities, the latter are excluded from the final dataset. In total, the dataset contains 2,023 municipalities[6], resulting in 46,534 municipality-year observations.

The aggregate firm-data are merged with location characteristics of each municipality. First, I use population for 1995 to 2017 originating from the INKAR database (BBSR, 2022a). Locally available knowledge is proxied by the number of universities in the municipality. Locations published in a public register of colleges and universities (Hochschulkompass, 2020) are used.[7] The data are time-variate albeit with 27 additional higher education institutions in 25 municipalities over the sample period.

4.3.3 Descriptive Statistics

Table 4.1 shows the distribution of digital firm birth for core cities vs. municipalities for the period from 1995 to 2017. There was an average of 53.1 firm births per core city, and 0.7 in municipalities. Berlin, had most firm births in absolute terms in 2015. Neuss, a municipality close to Dusseldorf with the highest firm birth registered 61 new firms in 2012. This pattern shows that digital firm birth is, in absolute terms, an urban phenomenon – the mean yields 0.05 per 1,000 inhabitants for municipalities and 0.11 for centers.

Table 4.1: Absolute Firm birth in centers and municipalities

	min	mean	max
Municipalities	0	0.7	61
Center	0	53.1	1507

Notes: The Table shows summary statistics on absolute firm birth in municipalities and the centers.

[4] Darmstadt/ Frankfurt/ Wiesbaden/ Mainz; Düsseldorf/ Duisburg/ Krefeld/ Mönchengladbach; Essen/ Bochum/ Dortmund/ Hagen; Köln/ Bonn; Ludwigshafen/ Mannheim; Nürnberg/ Erlangen; Berlin/ Potsdam.

[5] Augsburg, Bielefeld, Braunschweig, Bremen, Bremerhaven, Chemnitz, Dresden, Erfurt, Freiburg, Göttingen, Halle, Hamburg, Hannover, Heidelberg, Heilbronn, Hildesheim, Ingolstadt, Jena, Karlsruhe, Kassel, Kiel, Koblenz, Leipzig, Lübeck, Magdeburg, München, Münster, Oldenburg, Osnabrück, Paderborn, Pforzheim, Regensburg, Reutlingen, Rostock, Saarbrücken, Salzgitter, Siegen, Stuttgart, Trier, Ulm, Wolfsburg, Würzburg.

[6] Note that the final dataset does not cover all municipalities in Germany, but excludes municipalities not belonging to core-labor market regions. Core cities are excluded from the analysis.

[7] Addresses have been geocoded with the geocode command in R.

Figure 4.1 displays the pooled share of digital firm birth in centers and municipalities for the sample period. Hamburg, Berlin, Erfurt, Leipzig and Dresden show a high concentration of firm birth, as 73-88 % of all firms set up in the respective counties are in their respective core cities. West-German regions, the Ruhr-area in particular, show a more balanced distribution. This might be indicative of agglomeration benefits not decaying as strongly over distance. Taken together, this indicates that the spatial range of agglomeration benefits decreases faster around bigger cores, seeing that Hamburg and Berlin are the biggest cities in Germany. Further, these cities are more isolated in space that is they are very few neighbours that are big in size.

Table 4.2 shows summary statistics for the MURs and PURs (without the centers). While some monocentric cores were able to absorb 100 % of their entrepreneurship capital, the top city in a PUR could absorb 66 % (Cologne in 2001). This indicates different distribution channels of entrepreneurship capital for MURs and PURs. Presumably, municipalities in PURs can make use of a wider network of agglomeration externalities as laid out in Volgmann and Münter (2022). They argue that PURs have a lower mass to generate agglomeration externalities while cities serving as the single center of their region are typically more oriented towards their central business districts.

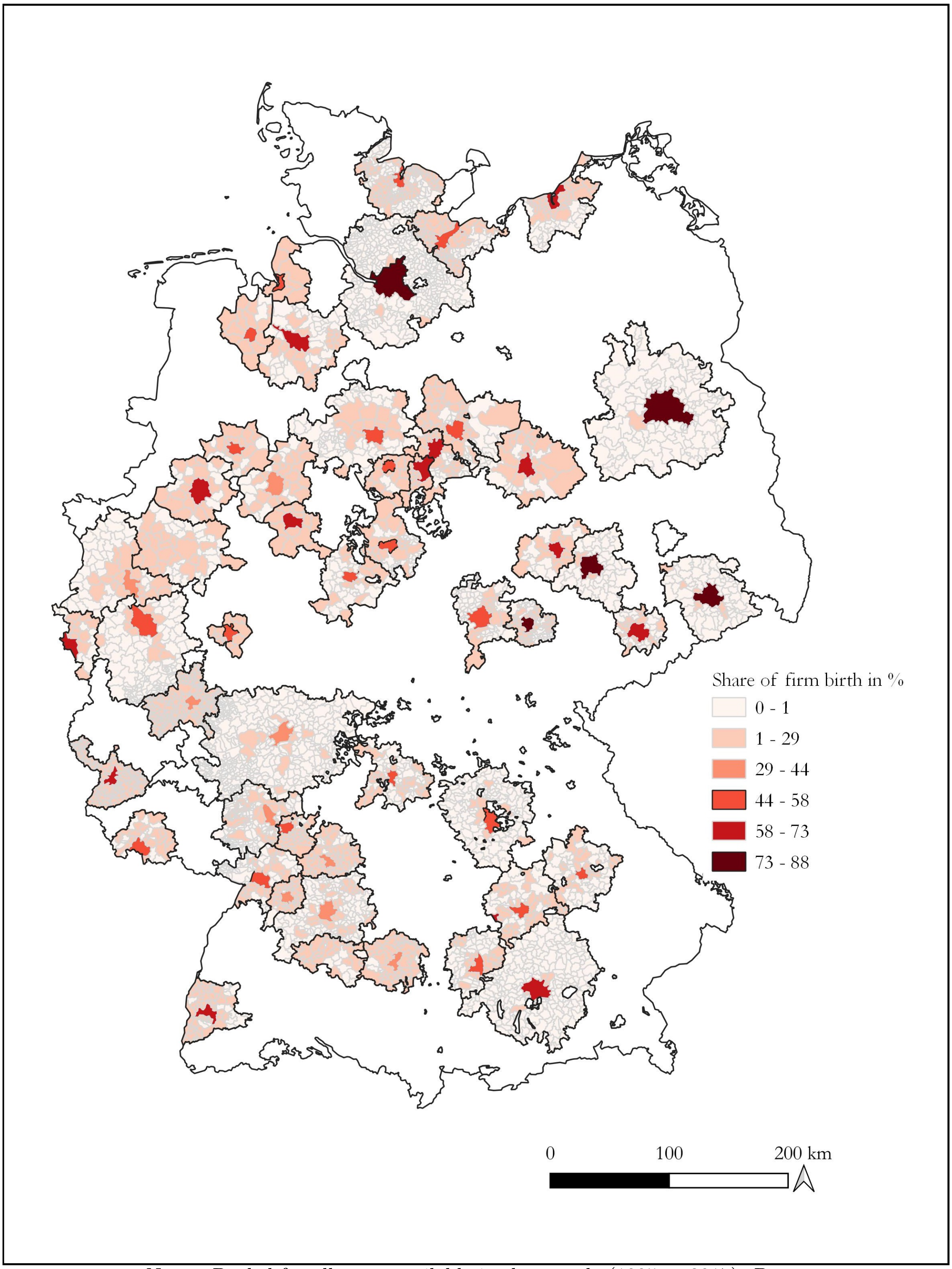
Table 4.2: Share of firm birth in regions including centers

	min	mean	max
MUR	0.0	1.18	100.0
PUR	0.0	1.03	60.2

Notes: The Table shows summary statistics on the share of firm birth in mono-and polycentric urban regions, including the centers.

Table A2.1 shows summary statistics of the control variables, Table A2.4 (see Appendix) shows a correlation table of the data. Especially the share of firm birth in the next center and the share of firm birth in municipalities have a correlation coefficient of only -0.003. However, the 'universities next center' and the 'population in the next center' show a rather high coefficient. Although multicollinearity between the variables should not be a major concern, additional regressions are presented in the robustness section.

Figure 4.1: Digital Firm Birth in Regions



Notes: Pooled for all years available in the sample (1995 to 2017). Data-source basemap: GeoBasis-DE / BKG 2021

4.4 Empirical Strategy

I use firm births in the digital sector as proxy for entrepreneurship capital as the variable reflects individuals capitalising on their novel ideas in a relatively low sunk cost environment. For setting up a digital business, capital costs for physical inputs are typically low. Hence, the location of digital firm births reflects a spatial knowledge allocation as the profit-maximizing firm locates within a labour market where (agglomeration) benefits outweigh the costs. Thus, firms reveal their preferences in terms of labour, amenities and market access by choosing a suitable location.

The dependent variable is the share of regional firm births in non-core municipalities which is defined as

$$Y_{i,t} = \frac{\sum k_{i,t}}{\sum k_{j,t}} * 100 \tag{4.1}$$

where profit maximizing firm k in year t is set up in municipality i. The sum of all firms set up in i relative to all firms set up in the same labour market region j reflects the municipalities' relative ability to attract firm birth.

The econometric analysis proceeds in two steps: First, I identify general determinants of local characteristics conducive to firm birth outside the administrative borders of the region's core cities. Second, I use two subsets of the original dataset with individual estimations for MURs and PURs to determine differences in the distribution of entrepreneurship capital for each morphological type.

All estimations are panel fixed effects models. With the inclusion of time – and region fixed effects, I control for region-specific time-invariant characteristics and yearly increments such as general trends in the industry or the economy at large. Additionally, all observable and unobservable effects which might vary on region- and time level are controlled for which reduces the threat of omitted variable bias.

The following Model (1) will be estimated using OLS with standard errors clustered on municipality level:

$$Y_{i,t} = \alpha + l_{i,t} + l_{c,t} + FB_{j,t} + SFB_{c,t} + d_{ic} + T_t + \gamma_{j,t} + \epsilon_{i,t} \tag{4.2}$$

The independent variables enter both models as follows: $l_{i,t}$ refers to locational characteristics in location i at time t possibly creating agglomeration externalities (population size divided by 10,000 as proxy for sharing; and Higher Education Institutions as proxy for learning). $l_{c,t}$ refers to the same locational characteristics, but in the center that is closest to the municipality as I assume the core to affect its labour markets' municipalities. $FB_{j,t}$ is the total firm birth in the region controlling for a general size-effect. $SFB_{c,t}$ is the share of firm birth in the center, d_{ic} is the physical distance between the geographic centers of the core city and the municipality. T_t is a year fixed effect controlling for common trends. $\gamma_{j,t}$ is the time-invariant region fixed effect capturing unobservable region-specific and time-invariant factors which are potentially correlated with the number of new firms. $\epsilon_{i,t}$ is the error term.

Examining differences between MURs and PURs, a second model will be estimated twice, one for each subset of the data (MURs and PURs separately). To specify the dynamics of agglomeration externalities more specifically, the baseline specification is modified as follows:

$$Y_{i,t} = \alpha + l_{i,t} + l_{c,t} + FB_{j,t} + +SF_{c,t} * dq_{ic} + T_t + \gamma_{j,t} + \epsilon_{i,t} \tag{4.3}$$

The predictors remain the same as in the baseline model, but the share of firm birth in the next center $SFB_{c,t}$ interacts with the distance. I use distance quintiles denoted as dq_{ic}. Each distance quintile is sensitive to the size of the individual labour market area. This measure was chosen over absolute distances because it captures the relative size of each region. The share of firm birth is expected to vary with the next center's share of firm birth in each quintile when agglomeration externalities diminish relative to the core size. Further, it smoothens the distribution of the distance measure because the fifth quintile would only be covered by very large regions.

4.5 Empirical Results

4.5.1 Local factors vs. core city characteristics

Table 4.3 presents the empirical results. Column (1) presents the baseline model that is the pooled, full-sample regression estimated by in Equation 4.2. This captures the first research question (Research Question 1) 'How do local factors and core city characteristics influence the intra-regional distribution of entrepreneurship capital?' The first ex-ante assumption was that general agglomeration externalities such as population (growth) and a local knowledge base create advantages over other municipalities in the labour market region.

Results show that population growth and being a university town are significant factors for materialising entrepreneurship capital. That is, for a population growth of 10,000 inhabitants, the share of firm birth increases by an average of 0.38 %. Additionally, hosting a university increases the share of firm birth by 0.95 %. Note that this is a fixed effect model and coefficients are interpreted as the changes in the parameters.

Table 4.3: Estimation Results Baseline Regressions

	(1)	(2)	(3)
Population: 10,000 inhabitants	0.380***	0.890***	0.210***
	(0.013)	(0.025)	(0.008)
University Dummy	0.950***	0.860***	0.280***
	(0.110)	(0.180)	(0.068)
Population: 10,000 inhabitants next center	0.008***	0.049***	0.0003
	(0.001)	(0.007)	(0.0003)
Universities next center	0.040***	0.020*	-0.002
	(0.005)	(0.011)	(0.003)
Absolute firm birth Region	-0.0002***	-0.001***	-0.0001
	(0.0001)	(0.0001)	(0.0001)
Share of firm birth next center	-0.039***	-0.054***	0.007***
	(0.003)	(0.004)	(0.001)
Distance to next center	-0.009***		
	(0.001)		
Distance Quintile2		-0.950***	-0.060***
		(0.210)	(0.021)
Distance Quintile3		-0.510	-0.080***
		(0.410)	(0.018)
Distance Quintile4		-0.320	-0.050***
		(0.330)	(0.019)
Distance Quintile5		-3.200***	-0.200***
		(0.790)	(0.024)
ShareFB nextCenter:Dist.Quintile2		0.012***	-0.007***
		(0.003)	(0.001)
ShareFB nextCenter:Dist.Quintile3		0.007	-0.007***
		(0.006)	(0.001)
ShareFB nextCenter:Dist.Quintile4		0.004	-0.007***
		(0.005)	(0.001)
ShareFB nextCenter:Dist.Quintile5		0.043***	-0.005***
		(0.013)	(0.001)
Time FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Observations	46,534	31,722	14,812
Adjusted R_2	0.170	0.190	0.380

Notes: Dependent variable is the share of firm birth. Column (1) refers to the baseline with the full dataset, Column (2) shows results for the subset of monocentric regions and Column (3) shows the subset of polycentric regions. Cluster Robust Standard Errors in parenthesis on municipality-level. Significant levels at ***p < 0.001; **p < 0.01; *p < 0.05

Population growth in the next center as well as an additional university do have a positive effect on the share of firm birth in municipalities. This indicates that firms in the regions' municipalities have access to additional agglomeration advantages provided by the respective core cities. This spillover effect may contribute to the general labour market. This finding provides evidence for a 'borrowing size' effect, where the municipalities benefit from their core (Phelps et al., 2001). This result is in line with Meijers et al. (2016), who also find individual growth to be more important than growth in 'connected' areas. Consistent with theoretical expectations (Delgado et al., 2010; Pijnenburg and Kholodilin, 2014), the results clearly show municipalities' own characteristics, such as population growth and locally available knowledge, are affecting the firm birth in a much stronger way than growth in the center, as indicated by the significantly larger coefficients.

However, growth in the digital industry, captured by absolute firm birth in the region, as well as the share of firm birth in the next center are negatively related to the share of firm birth in the municipality. These results show that non-core municipalities lose out in terms of relative attractiveness to the core's growing industry. Presumably, growth in the core lowers the costs of starting a business while simultaneously increasing learning effects within the industry. These learning effects work on very small scales within a few neighbourhood blocks (Arzaghi and Henderson, 2008) and contribute to outweighing agglomeration costs.

Altogether, 'borrowing size' effects and 'agglomeration shadows' seem to co-exist along different channels: For generic population growth and institutional knowledge in Higher Education Institutions I find a borrowing size effect while specific industry-relevant factors such as within industry spillovers contribute to a stronger concentration of firms within core cities. Note that this finding does not necessarily contradict Pijnenburg and Kholodilin (2014) and Fossen and Martin (2018) findings that neighbours profit from core growth in absolute terms.

Lastly, the closer a municipality is located to the center, the larger the share of firm birth that proxies entrepreneurship capital. The share of firm birth decays by approximately 0.01 % with each additional kilometer of distance to the urban core. That confirms the hypothesis that agglomeration advantages decrease with distance and a significant advantage of small municipalities' lies in their geographical location, i.e. their proximity to the next core.

4.5.2 MURs vs. PURs

Columns (2) and (3) in Table 4.3 provide estimation results for the second research question 'Does the intra-regional distribution of entrepreneurship capital differ in MURs and PURs?' Results show that MURs and PURs indeed differ in the distribution of entrepreneurship capital.

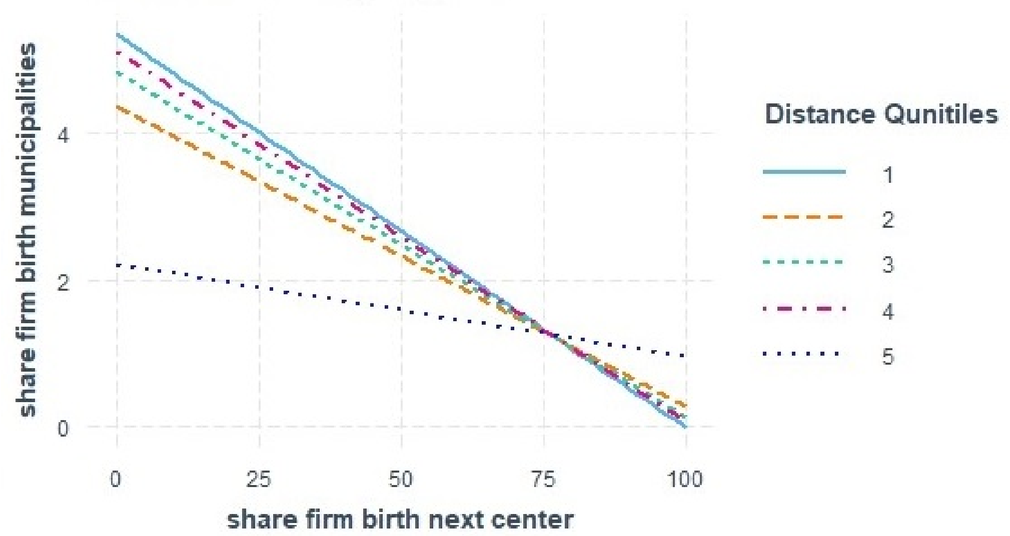
In MURs, 10,000 additional inhabitants increases the share of firm birth for municipalities by 0.89 %. For PURs, there is only an increase of 0.21 % for the same amount of population growth. Results show positive effects if the municipalities host a university for both morphological types. However, a university within a MUR has a three times lager impact. Therefore, both population growth - which can be interpreted as sharing externalities - as well as institutionalized knowledge boost relative advantages of municipalities in a MUR four times as much as for municipalities in PURs. Concerning the interdependence of the municipalities' entrepreneurship capital with the economic cores of the regions, MURs and PURs show significant differences. In MURs, population growth in the next center increases the share of firm birth in its surrounding areas. Similar population growth in centers within a PUR has no significant effect on its' municipalities, that is their shared labour market. Put differently, PURs do not create significant additional sharing externalities by means of population growth. Further, an additional university in a MURs' core city significantly increases the relative distribution of entrepreneurship capital in municipalities while there is no such effect in PURs.

These differences is also visible for general industry growth: A monocentric core absorbs absolute industry growth by lowering the municipalities' shares of firm birth, but general industry growth does not impact the PUR's share. Overall, this shows that municipalities in MURs are more dependent on the core's general as well as industry-specific development than municipalities in PURs, as the latter presumably access agglomeration externalities from several (at least two) cores.

Figures 4.2 and 4.3 visualize the relationship of distance and the next center's share of firm birth on the share of firm birth in the municipality. Note that the y-axes of the figures differ and effects in MURs are significantly larger in size than for PURs as polycenters are naturally more dispersed. The coefficient of the interaction term shows the effect of $x1$ on y for a given quintile. That is, at certain distances the share of firm birth in the next center has different effects on y. If the interaction coefficient is positive the effect of $x1$ on y increases as $x2$ increases, if negative it decreases. In other words, I expect the distribution of firm birth to wane with increasing distance from the core.

For the MUR, Figure 4.2 shows a steep negative slope of the interaction for the first distance quantile (municipalities located closest to the center). The higher the share of firm birth in the center, the lower the share of firm birth in the municipality with a mean distance of 13 km. For the second quintile - municipalities not being located directly next to the center (24 km mean distance) - the slope is less steep than for the first quintile. The fifth quintile covering municipalities on the outer bounds of the labour market region are least dependent on the dynamics in the inner core. Overall, the results show core cities in MURs are clearly the centers of economic activity and strongly interact with their direct neighbours, as the slope of the first quintile is the steepest. Nevertheless, this result does not provide an answer to whether there is an 'agglomeration shadow' or a 'borrowing size' effect; nonetheless the findings indicate that municipalities' dependence on the core is strongest for direct neighbours in the first quintile either way.

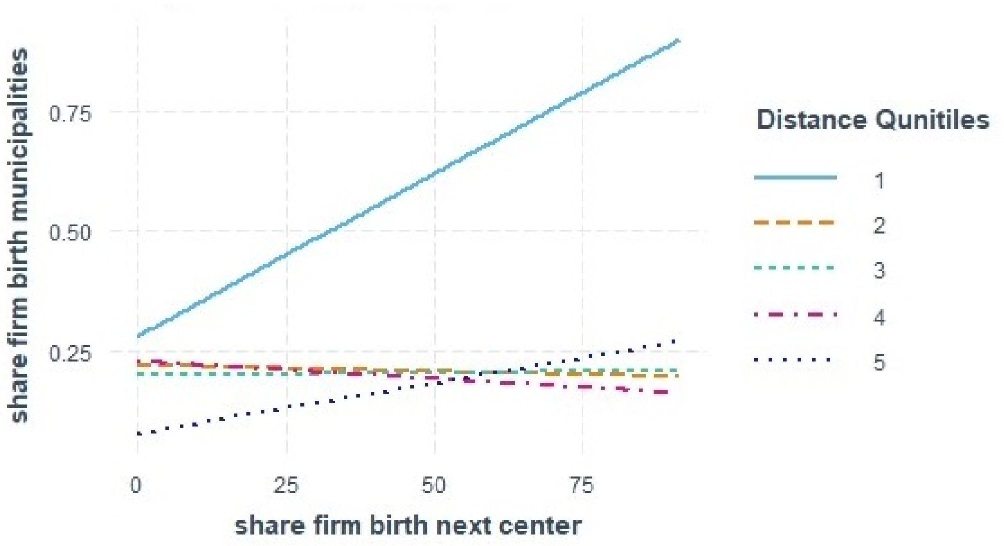
Figure 4.2: Post-estimation Interaction Plot of MURs



Notes: Note that Quintile 3 and 4 are insignificant in the estimation results.

In PURs (see Figure 4.3), the distribution of entrepreneurship capital is notably different from MURs. For the first distance quintile, the effect is opposite to that for monocentric regions: An increase in the share of firm birth in the next center also increases the share of firm birth in its direct neighbours. An increase in the relative advantage of the next core city therefore increases the relative advantages of its neighbours. Interestingly, this effect is detected for the first quintile, so direct neighbours only. Therefore, the spatial spillover of increasing industry concentration effects is limited to the first distance quintile around a core city in a PUR. For the second, third and fourth distance quintile there are very small negative effects when the entrepreneurship capital in the core city increases. Despite that, for regions in the fifth distance quintile, the effect turns positive again. However, this finding has to be taken with a grain of salt, as being located within the fifth distance quintile might imply a location between two core cities. Possibly, the outcome reflects an ‘overlap’ of agglomeration advantages in the outer-bounds of the labour market where costs of agglomeration are relatively high compared to the gains.

Figure 4.3: Post-estimation Interaction Plot of PURs



Notes:all coefficients are significant in the estimation.

Overall, these results show that the intra-regional distribution of entrepreneurship capital differ in MURs and PURs as the regional distribution of entrepreneurship capital is shaped by the urban core in MURs, but growing municipalities with better knowledge infrastructure have a relative advantage to their sparring partners. In sum, municipalities in MURs benefit from universities’ sharing and general knowledge, more generally perhaps also from typical urbanisation externalities. Rosenthal and Strange (2003) argue that urbanisation effects reflect the trade-off between benefits in dense areas and congestion costs. An increase in the industry-density in the core that is localisation externalities, on average, results in the core throwing an agglomeration shadow. Ceteris paribus, an above average growth in the digital industry in the core absorbs its’ small neighbours entrepreneurship capital. Thereby, increasing advantages from e.g. tacit knowledge exchange exceed congestion costs. This is in line with the literature, where localisation externalities are usually found to decay more rapidly with distance than urbanisation externalities (Rosenthal and Strange, 2003; Arzaghi and Henderson, 2008; Andersson et al., 2019).

Municipalities in PURs, however, are less affected by changes in urbanisation externalities, specifically in the next core. PURs probably offer more evenly distributed agglomeration externalities, as firms in municipalities can access several cores (Volgmann and Rusche, 2020). Thus, they are less dependent on urbanisation externalities of one particular neighbouring center.

Nevertheless, for localisation, the results show a reverse effect for municipalities being located closest to urban cores in terms of industry-growth. This hints towards the existence of competition effects between urban cores in PURs. Thereby, more competition attracts rather than repels firms. This is in line Kim et al. (2022) who analyse the patterns of firm formation of 242 industries in 508 regions over 15 years in Australia. While this seems at odds with the literature presented above, it still shows that digital firms gain advantages from being located in clusters (Delgado et al., 2010) and close to industry-specific knowledge, as the cores gain relative attractiveness over their competition cores within the labour market region.

Robustness

To analyse and establish the robustness of the results, several tests have been conducted. Providing a more detailed picture of the importance of the distance, the empirical models were estimated again without the interaction term plus a continuous distance measure for MURs and PURs (Table A2.5). Results for the continuous measures show a stronger spatial decay of the share of firm birth in MURs compared to PURs (see Columns (2) and (3) in A2.5). Leaving out the interaction term as presented in the baseline model, results show the decay of the share of firm birth more precisely. Effects for the quintiles decay continuously, but the coefficient for the second quintile is larger than for the third quintile in MURs. This is indicative for agglomeration shadows diminishing over distance. That is, municipalities in the second quintile are overshadowed by the first quintile, but this shadow seems to fade out for the third and fourth quintile. Municipalities in the fifth quintile however are too far off to reach out to externalities provided by the respective core city.

The sensitivity of the results towards the chosen quantiles was tested by employing the absolute measure of quantiles. In contrast to the relative measure as described above, the quantiles are benchmarked against the absolute distance of the largest labour market. The regressions (without interactions) have been repeated. Table A2.6 presents the results, Figure

A2.1 illustrates the coefficients and the distribution of the standard errors. Figure A2.2 shows the coefficients of the relative distance quintiles for comparison. Note that the x-axis scales are different. To provide a better comparison, Figure A2.3 presents Quintiles 2-4 separately. First, results are strongly driven by the definition of the quintiles and relative measures produce less differences in MURs and PURs than the absolute measure.

The baseline regression including interaction effects was repeated with the alternative quintile measure. Figures A2.4 and A2.5 show post-estimation plots similar the main section. The effects point to the same direction as the baseline model. However, especially for PURs, the results differ in the second to forth quintile, as they show a positive slope. This underlines the overall picture, that municipalities in PURs are likely to gain advantages from 'borrowing size'.

Table A2.7 presents a variant of the baseline model, estimating of the full sample while including a dummy variable for MURs. Results show that municipalities in MURs generate significantly more share of firm births as municipalities in PURs. This underlines the wider dispersion of the industry in PURs.

As multicollinearity might be a concern for the relations of the municipalities and the centers (as the correlation coefficient was high, see Table A2.8), the regressions have been conducted with the municipality's characteristics only. Results presented in Table A8 show that the results for individual municipalities are robust.

To ensure the suitability of the share of firm birth as a measure, the regressions have been performed with the absolute firm birth as dependent variables (see Table A2.9). Results show that the municipalities individual factors are less pronounced in absolute terms. That means, population growth and the presence of a university have similar effects in MURs and PURs. However, results show a similar pattern as for the relative measure for the characteristics in the next core cities: Municipalities in MURs take advantages of universities in the next core, for municipalities in PURs we see an opposite effect. Nevertheless, a general growth in the industry in the region has a stronger effect for MUR than PUR municipalities.

4.6 Conclusions

This Chapter analyses regional firm birth patterns of the digital industry apart from core cities to shed light on the spatial distribution of entrepreneurship capital that determines the general development of regions in the long run and the dependence of municipalities on their core cities. The municipalities provide labour-inflows into the core, while the core itself provides agglomeration externalities that contribute to higher productivity and higher wages. Nevertheless, it is in the interest of municipalities' political decision makers to attract some of the entrepreneurship capital of their regions as they gain attractiveness and income flows through business taxes by hosting firms. Further, the Chapter considers a link between the intellectual baselines of 'agglomeration shadows' and 'borrowing size' and agglomeration economics. Therefore, the first research question is 'How do local factors and core city characteristics influence the intra-regional distribution of entrepreneurship capital?'

The empirical study shows that municipalities' digital industry development depends on individual characteristics as well as those from the next respective core city: For generic population growth and institutional knowledge in HEI, I find a positive effect for individual characteristics while specific industry-relevant factors such as within industry spillovers contribute to a stronger concentration of firms within core cities. Results further show that relative advantages of small municipalities next to 'giant neighbours' strongly depend on the distance between the dwarf and the giants. The second research question is: 'Does the intra-regional distribution of entrepreneurship capital differ in MURs and PURs?' The analysis shows that MURs and PURs differ in the general morphology of industry dynamics. This puts the results from the first research question into perspective, as the industry development mechanisms differ.

Results show that MURs tend to absorb entrepreneurship capital from their direct neighbours with growing industries, population growth and strong academic landscapes. Thus, MURs with one successful digital center lay an agglomeration shadow over their direct neighbours in terms of digital start-ups. Successful cores in PURs, however, serve their neighbours by increasing their relative attractiveness over other municipalities and allow them to stand on giant shoulders while borrowing industry-specific externalities. This gain in PURs is specific to the industry, as population growth and institutionalized knowledge in additional universities are not significant for PURs. On top of that, this indicates competition effects of the individual centers within PURs.

This result reveals important policy implications. The success of the same policy intervention in two similar municipalities may have different impacts depending on the relative location of the municipality in its specific labour market. The EU states that the support of polycentric development can create a critical economic mass by combining efforts of urban centers. An understanding of joint competitive advantages can help cooperating cities to strengthen their competitive resourcing power in a greater regional context (ESPON, 2017). The results of the chapter show that a specialisation of individual cities in a PURs could be a promising approach.

All in all, the results are likely to be transferable to other primarily European contexts. PURs also occur in other European countries, prominent examples being the Milan-Bergamo region or Randstad in the Netherlands, which includes Amsterdam, Rotterdam and The Hague. The German context primarily offers the advantage of sufficient sample size of both MURs and PURs.

Chapter 5

Universal University? Micro-geographical Assessments of External Knowledge Inputs for Digital Firms from a Spatial Perspective

Abstract

This Chapter links clustering of digital firms and proximity to knowledge institutions, i.e. research institutes and HEIs, on neighbourhood level. Using a new micro-geographical dataset on 24,614 firm births in the digital sector from 2008 to 2016 in Berlin, Hamburg and Munich, I investigate whether firm birth occurs close to distinct institutional knowledge providers. The panel fixed effects regression models show significant firm clustering within cities close to knowledge institutions that provide tacit industry-related knowledge – data science and design skills. Further, the more tacit knowledge a knowledge institution contains, the closer firms locate within cities.

5.1 Introduction

For young knowledge intensive firms, access to external knowledge sources is essential for competitiveness. Similarly, incumbent firms and knowledge institutions are drivers of entrepreneurship as new ideas and knowledge spillovers are channeled into new firms (Acs et al., 2009). Thus, knowledge's geographical accessibility is of major interest for urban policy makers as it has important implications on where to intervene with tax money to spur on growth by means of innovation or industry development. The digital industry in particular has been of major interest for local policy makers, as it complements many, almost all other sectors, contributes significantly to the GDP, demands high skilled and well payed employees and offers few negative externalities in terms of land use and emissions.

Tacit knowledge spillovers, a crucial input for knowledge-intensive firms, work on small neighbourhood-scales within industries (Arzaghi and Henderson, 2008). Little is known about geographical proximity in the dissemination of knowledge between HEIs and research institutes to new firm formation, although being the theoretical backbone of endogenous growth models (Romer, 1990). Nevertheless, over the past decades, universities have been required to open the door to the 'ivory towers' and play the roles of regional powerhouses of knowledge transfer to foster entrepreneurship and growth (Geuna and Muscio, 2009; Ghinamo, 2012).

This chapter links spatial proximity between firm birth activity in the digital sector and knowledge institutions within cities. I geocode and aggregate 24,614 firm births between 2008 and 2016 in three cities on 1x1 km^2 grids. I complement these new firm data with a tailor-made granular dataset on HEIs and research institutes. With this empirical set-up, I test whether there is a higher density of ICT start-ups clustering around distinct knowledge institutions while controlling for common location factors.

The cities of Berlin, Hamburg and Munich are of particular interest as showing the highest start-up dynamics in Germany, with the fourth biggest ICT location Frankfurt/Main only producing 23 % of Berlin's 1,423 of startups (in 2016). Berlins ICT sector employs 130,900 people (36 per 1,000 inhabitants) while Hamburg employs 70,200 (38 per 1,000 inhabitants) and Munich 92,770 (63 per 1,000 inhabitants) (Bundesagentur für Arbeit, 2021; City of Munich, 2022). Thus, Germany and in particular these three cities with their tight knowledge infrastructure as well as spatially disaggregated information on digital firms offers an excellent opportunity to study knowledge flows and their importance in fostering start-up ecosystems in the ICT sector.

I explore the knowledge ecosystem in three ways: First, I test whether there is quantifiable, significant firm clustering close to HEIs and research institutes as this is not entirely clear in the literature. For example, Rammer et al. (2020) find innovative startups in Berlin locate significantly closer to universities while Duvivier et al. (2018) cannot find a similar effect for ‘new economy’ employment in Canadian cities when controlling for other location factors.

Second, advantages from localisation (same industry spillover) work on smaller scales than advantages from urbanisation (diverse spillover, e.g. Andersson et al. (2019)). It is almost entirely unclear whether the HEI-industry knowledge exchange requires ‘economic proximity’, that is similarity of knowledge contents. Therefore, I characterise universities by their departments and the research area of research institutes and classify them as related and unrelated to ICT firms. Thus, I test whether firm clustering occurs close to any knowledge content or if specific inputs are required.

Third, the HEI landscape in Germany consists of research universities, universities of applied sciences (technical colleges, UAS) and universities of music, arts and design (hereafter design universities) that differ substantially in their setups. Based on differentiated knowledge bases (Grillitsch et al., 2017), I proxy highly codified analytical knowledge that can be transferred over long distances by research universities. UAS rather host more engineering-based, problem solving, applied synthetic knowledge that is not transferred over long distances as easily as analytical knowledge. Finally, symbolic, highly context specific knowledge requires face-to-face contacts when transmitted and is mostly found in design universities. I test whether firm clustering occurs in accordance to the knowledge bases’ transferability on micro-scales leaving important takeaways for urban policies.

I find significant firm clustering in neighbourhoods with HEIs and research institutes. However, cluster effects within the industry decay more rapidly over distance (88 %) than for HEIs (54 %) and research institutes (59 %).[1] This indicates tacit spillovers from knowledge institutions to be conducive to new firms. At the same time, transmission channels less sensitive to geography, (e.g. networks) seem to be stronger than for within-industry linkages. This finding provides empirical evidence for Duvivier et al.’s (2018) suggestion that university-industry spillovers benefit many firms within the city.

I further find a significant effect on neighbourhood-level firm birth for HEIs specialized in IT and data science, research institutes for social sciences and design departments. This is reflective of digital firm’s desire to be closest to a knowledge stock that can directly be transferred into

[1] That is the effect of the number of firms within the same neighbourhood is 88 % higher than for contiguous neighbourhoods.

products. These results are similar to Andersson et al. (2019) and Lavoratori and Castellani (2021) findings for within-industry spillovers, however, their results have not been investigated for university-industry links on such small spatial scales before.

Finally, results indicate that the transferability of knowledge bases work on much shorter distances as the to-date findings that distinguish local, regional and national levels: I find no significant firm birth close to research universities, while there are strong positive effects for UAS and design universities. This novel finding is of great relevance for policy makers as universities in particular are an often-used outlet for public funding to accelerate firm birth activity. In Germany, this is often done by subsidizing office spaces close to (research) universities, for example. This Chapter, however, shows that a fine-graded political funding and presumably an industry-specific location policy vis-a-vis co-location to a HEI are needed.

5.2 Theoretical Background, Related Literature and Hypothesis

5.2.1 Micro-Agglomeration, localisation and urbanisation and HEI

Knowledge spillovers are main drivers of local innovation (Rosenthal and Strange, 2003). If the expected value of a new piece of knowledge or idea is higher for an individual than for a decision maker within an incumbent firm or university, the individual will start a new business and become an entrepreneur (Acs et al., 2009; Audretsch et al., 2008). Accordingly, the presence of incumbent firms and knowledge institutions can result in new firm formation, manifesting the local knowledge base. Nevertheless, knowledge spillovers are equally crucial for existing firms: For innovating firms, constraints in technical capabilities outside of their existing knowledge are likely to cause problems (Schartinger et al., 2001). Because of high costs for internalizing knowledge, firms rely on retrieving outside knowledge. This spillovers have been shown to work on small spatial scales within cities, which is why being close to knowledge sources is crucial for small firms (Van Soest et al., 2006; Larsson, 2014; Jang et al., 2017; Rammer et al., 2020; Roche, 2020). In this context, the scholarly debate also reflects what kind of knowledge spillovers and business-environments are conducive to firm development and competitiveness. Recent studies on firm-to-firm externalities find that localisation - similar, same industry knowledge (Porter, 1990) and urbanisation (diverse, other industries-knowledge; Jacobs (1969)) - play a vital role for firm productivity.

However, these effects seem to operate on different spatial scales: localisation externalities operate in neighbourhoods of one square kilometer and less which is why similar firms cluster in that radius (Andersson et al., 2019; Lavoratori and Castellani, 2021). In contrast, benefits incurring from urbanisation operate on neighbourhood- and city-level. Especially for high-tech and knowledge-intensive firm births, the potential benefits of 'cross-fertilisation' between industries to generate new ideas and innovation stand out (Andersson et al., 2019).

Besides industry-knowledge, knowledge institutions generate externalities. Rammer et al. (2020) identify proximity to research institutes and universities as a distinctive feature for the location choice of innovative firms by using a matching approach on panel data covering 2011 to 2015 with 3,723 firms in Berlin. On the contrary, an investigation of ICT-employment density using microdata for three Canadian cities, does not show a significant effect for research universities when controlling for other location factors (Duvivier et al., 2018). The authors conclude that university spillovers must be relevant for all locations within the cities.

The key question in this Chapter is whether digital firms derive advantages from being close to knowledge institutions within the city. Good et al. (2019) and Fabiano et al. (2020) show manifold university-industry linkages used for public-private knowledge transfers in their literature reviews on technology transfer ecosystems in academia. Especially tacit university-industry spillovers via incubators require close proximity (Fabiano et al., 2020). Kerr and Kominers (2015) point out individual channels working on the regional level, such as by the labour market. Therefore, firms that locate close to knowledge institutions should gain some 'top-up' advantage that is not transferable by other channels. Based on these considerations, Hypothesis 1 is that clusters of digital firms occur in close proximity to HEIs within cities.

5.2.2 Knowledge content and input for ICT firms

In the light of localisation vs. urbanisation and university-industry linkages, it is unclear whether university-knowledge-inputs are a 'localisation' or 'urbanisation' benefit for digital firms. In an industry-context, Frenken et al. (2007) argue that not a generic diversity, but a local variety of related firms offers know-how for knowledge transfers. Mainly, spillovers require at least some similarity concerning knowledge bases, competencies or skills, labour pools and technologies to ensure absorptive capacity (Boschma, 2005). Therefore, the potential of inter-firm knowledge spillovers is higher for firms operating in similar industries, that is, they share 'economic proximity' (Van Oort et al., 2015).

Transferring this into an HEI-industry setup, this Chapter asks what tacit knowledge contents digital firms gain from HEIs and research institutes vis-a-vis the localisation economies gained from similar firms nearby. This has not been investigated within cities.[2] Especially for the digital sector being dominated by small firms (more than 90 % of all ICT firms had less than 10 employees in 2017; Destatis (2021)), collaboration between public and private sector is of particular importance (Cornett, 2009). Finding 'best matches' helps efficient city planning and providing optimal locations as breeding grounds for new and emerging firms.

Notwithstanding the heterogeneity of the sector itself, digital companies combining several areas of knowledge enhance competitiveness. First, a digital firm needs knowledge on the fundamentals of its business model that is technological skills (as implied by e.g. Grillitsch et al. (2019)). Second, business knowledge such as accounting and sales is needed. Third, scholars increasingly acknowledge the role of aesthetics and design for innovation (Secundo et al., 2020). Tödtling and Grillitsch (2015) find firms with internal competencies in design and product or process management to be more innovative. Therefore, technological knowledge or data science, business-knowledge and design competencies are defined as a related variety for digital firms. I hyposize that digital firms cluster in neighbourhoods where related knowledge can be retrieved from HEIs and research institutes (Hypothesis 2).

5.2.3 HEI types and differentiated knowledge bases

Knowledge creation and transmission comes in different shapes. Agglomeration theory shows transmission of tacit knowledge via face-to-face contacts to be key for understanding inner-city clustering. Other scholars argue for the existence of 'knowledge bases' going beyond traditional tacit and codified knowledge typologies (Plum and Hassink, 2011). Knowledge bases are categorized in analytical (science-based), synthetic (engineering-based) and symbolic (creativity-based) knowledge (Asheim and Gertler, 2005). Each base represents different combinations of codified and tacit knowledge, knowledge sources, interactions and transmission channels (see Table 5.1 for an overview). For the purpose of this Chapter, the taxonomy of the knowledge-base is applied and empirically tested using the various forms of HEIs in Germany as proxies for each base.

Analytical knowledge refers to scientific, theoretical knowledge based on formal models (Asheim et al., 2011). It is predominantly embodied in basic research and codified in scientific articles, reports and patents – and thus mobile and transferable over long distances (Moodys-

[2] Fritsch and Aamoucke (2017) link applied and natural science to innovative startups in regions. However, these results cannot be transferred to within city dynamics.

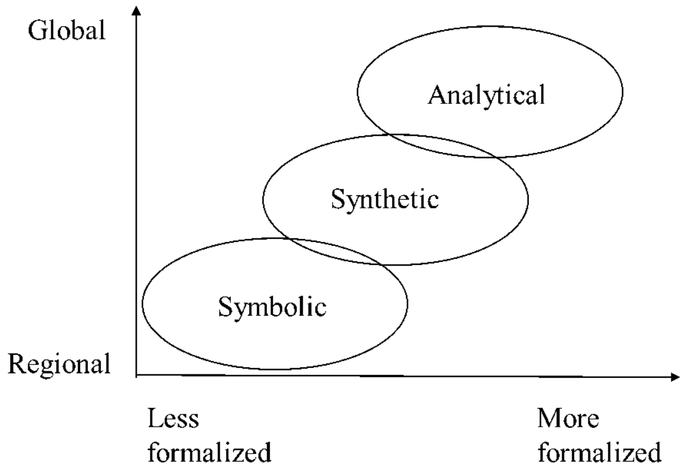
Table 5.1: Typology of the Knowledge Bases: Differentiated Knowledge Base Approach

	Analytical (science based)	Symbolic (engineering based)	Synthetic (design based)
Rationale for knowledge creation	Developing new knowledge about natural systems by applying scientific laws; know-why	Applying or combining existing knowledge in new ways; know-how	Creating meaning, desire, aesthetic qualities, affect, intangibles, symbols; images; know-who
Development and use of knowledge	Scientific knowledge, models, deductive	Problem solving, custom production, inductive	Creative process
Purpose of knowledge creation	Theoretically understanding natural or social systems, confirming or rejecting dominant scientific laws or defining new ones	Designing or constructing instrumental solutions to specific human problems	Creating socio-cultural meanings and interpretations of artefacts and their use
Typical target of innovation	Improvement of cognitive/theoretical models of products, processes or organizations	Change of functional attributes of products, processes or organizations	Change of aesthetic and semiotic features of products, processes or organizations
Actors involved	Collaboration within and between research units	Interactive learning with customers and suppliers	Experimentation in studios and project teams
Knowledge types	Strong codified knowledge content, highly abstract, universal	Partially codified knowledge, strong tacit component, more context specific	Importance of interpretation, creativity, cultural knowledge, sign value, implies strong context specificity
Type of knowledge created	Mainly codified, highly abstract and universal knowledge	Mainly tacit, context specific practical knowledge but important codified component	Strongly tacit, context-specific, semiotic context
Importance of spatial proximity	Meaning relatively constant between places	Meaning varies substantially between places	Meaning highly variable between place, class and gender

Notes: Own Modification after Asheim et al. (2011) and Manniche (2012).

son, 2008; Asheim et al., 2011) (see Figure 5.1 for a visualisation in the spatial characteristics). Research universities (Universitäten) focus on imparting theoretical knowledge and methodological competence, which is why their inherent knowledge is predominantly scientific and research-oriented (Hochschulkompass 2022). Therefore, research universities proxy knowledge sources for analytical knowledge (Tödtling and Grillitsch, 2015).

Figure 5.1: Codification and Geography of Knowledge Bases



Notes: The figure visualises the knowledge bases along the lines of the spatial level of transferability.

Synthetic knowledge refers to applied knowledge with problem-solving and engineering skills. Innovations mainly take place through the application or combination of existing knowledge. Such engineering work is partly codified (e.g. technical blueprints). It is the result of interactive learning and learning-by-doing. Thus, tacit knowledge is of greater importance for synthetic than for analytical knowledge building and therefore more sensitive to proximity (Asheim and Hansen, 2009; Plum and Hassink, 2011). UAS are characterised by a strong practical application of knowledge. The range of subjects is predominantly not as extensive as at research universities and is mostly concentrated on technical engineering, economics and social sciences (Hochschulkompass, 2022). This is why UAS proxy synthetic knowledge (Tödtling and Grillitsch, 2015).

The symbolic knowledge base accounts for design and aesthetics in products and services. Specialized abilities are interpretation and creativity rather than plain information processing (Martin and Moodysson, 2013). Due to the high cultural embeddedness, symbolic knowledge is characterised by a strong tacit component and usually highly context specific (Asheim et al., 2011). Therefore, it is the most sensitive to geography of all knowledge bases. Cross-fertilisation between professionals and sectors largely depends on informal interactions and buzz in non-commercial, daily-live contextual settings like street cultures or public events (Manniche, 2012). Universities for arts, music, and design offer students an academic education in the visual, creative and performing arts as well as in musical subjects. They are characterised by the special relationship between artistic, pedagogical and scientific education (Hochschulkompass, 2022). Therefore, design universities represent symbolic knowledge.

For the ICT sector, Trippl et al. (2009) find that firms combine knowledge from all three bases at the local level. Thus, firms are expected to locate closer to symbolic (design universities) than analytical (research universities) and synthetic (UAS) knowledge bases (Hypothesis 3).

5.3 Data

The tailor-made dataset encompasses three main components. First, firm-level data on young digital firms is used. Second, I use a rich knowledge infrastructure database on research institutes and HEIs. Third, the dataset includes variables for economic activity, socio-demographic conditions and infrastructure. This enables linking economic activity on neighbourhood level to neighbourhood characteristics and knowledge sources within cities to disentangle cluster effects and the underlying mechanisms.

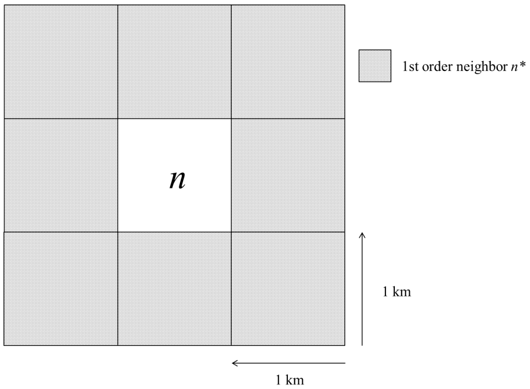
5.3.1 Grid-level analysis

The key advantage of using grids is that the position of the squares is independent from economic activity and thus tackling issues of endogeneity, while allowing to investigate externalities and spillover effects on small spatial scales. Following Andersson et al. (2019), the latent "true" scale of the mentioned externalities is unclear and could possibly cover several squares. To address this issue, this empirical analysis employs two spatial scales, the neighbourhood[3] (1x1 km^2 grid indicated by n) as well as its first-order neighbours (3x3 km^2 grid indicated by n^*). By that, each grid (n) has eight neighbour grids (n^*). This allows to test for possible decays of the expected effects within cities (see Figure 5.2).

The three cities analysed in this Chapter are Berlin (920 grids), Hamburg (795 grids) and Munich (350 grids). As the three biggest cities in Germany, they all host a thick knowledge base with several HEIs and research institutes. Although they differ in legislative characteristics

[3]Neighbourhood refers to the spatial level and does not cover a 'functional' neighbourhood. The grids refer to standardized EU-INSPIRE Grids.

Figure 5.2: Neighbourhood and Grid-Level



Notes: The figure visualises the neighbourhoods and the first-order neighbors.

(Berlin is the capital of Germany and federal state, Hamburg is a federal state on its own while Munich is the capital city of the federal state Bavaria), they all are considered the economic powerhouse of their regions.

5.3.2 Geo-coded firm level panel dataset

The uniqueness of the firm-data lies in the precise point-level tracking of individual firm locations. The analysis covers companies which entered the market between 2008 and 2016.[4] The data originate from statutory publications of German corporations and is provided by North Data (2019). Firm information includes date of incorporation, date of termination (if applicable), economic field, a description of the company's main business area and address history. The data does not include individual firm information such as financials or the number of employees.

As there is no agreed-upon definition of the digital economy, for the purpose of this Chapter, a digital firm is defined as information-technology driven and internet-based. I selected firms using NACE codes: General programming activities, software development, web portals, data processing, and the development of web pages, processing, hosting and related activities and web portals[5] (Weber et al., 2018). Yet, standard industry classification systems have limitations, especially regarding industries that cross-over traditional product categories like the digital sector (Oakey et al., 2001). Since digital business models complement many other sectors, firms may be registered in other NACE codes despite running a digital business model. Therefore the result-

[4]Due to data-availability of other location characteristics, the analysis covers eight years (2009 to 2016).
[5]62.01.0, 62.01.1, 62.01.9, 62.02.0, 62.03.0, 62.09.0, 63.11.0, 63.12.0

ing sample contains firms characterised by the core knowledge on which their competitiveness ultimately draws.

Identifying these firms, the description of the company's main business area is used. By applying a word-search selection, firms not registered in the ICT sector but operating on a digital business model were added to the dataset.[6] The resulting dataset comprises 24,614 individual firms with a total of 101,721 firm-year observations. By this, the data partly capture a related variety and economic proximity, based on shared knowledge bases beyond classical industry classifications.

The resulting panel dataset consists of firm-year observations. A firm's location in a given year is the location as of 31 December. The data only includes the headquarters, possible subsidiaries are not considered. Establishments exiting the market are excluded from the panel dataset after the year of deletion in the register. The location is available on point-level but is aggregated to the neighbourhood/1x1km^2 grid level (n) resulting in the key measure of the number of firms per grid. In total, the dataset contains 16,520 grid-year observations for the three cities.

5.3.3 Higher Education Institutions and research institutes

To explore the role of knowledge institutions conducive to ICT clustering, multi-tier data on HEIs and research institutes has been collected.

The disaggregated dataset contains the point-location of departments (Fakultäten) of HEIs. I distinguish departments of economics and social sciences; health and law; science, technology, engineering and mathematics Science, Technology, Engineering, and Mathematics (STEM); computer science; and departments for arts, music and design. For HEIs spreading over several locations, the exact location of the department is used. The data originates from Hochschulkompass (2020) while the location of departments has been drawn from the HEIs' websites. The number of students is not covered because it is not reliably available over time and locations. Further, the data allows distinguishing between research universities, UAS and universities for music, arts and design, covering a total of 76 HEIs in 108 unique locations.

For research institutes, the locations of institutes belonging to the four major German re-

[6]First, the description of the identified ICT firms is used to retrieve the most frequently used words related to IT and software. By these key-words, firms operating on digital business models. I exclude firms that only distribute their products via a webpage. For example, a firm registered in „Placement of workers“ is included, with explicitly stating to operate on a social networking platform for skills enhancement and marketing and to offer Internet-based services. A full list of keywords and in/exclusion of firms can be obtained from the author by request. (Note that keywords are in German).

search associations (Fraunhofer Institut, 2019; Helmholz Gesellschaft, 2019; Leibniz Association, 2019; Max-Planck-Institute, 2019), and institutes funded by the federal states as well as the national government (Forschungseinrichtungen des Bundes und der Länder, (OEFW, 2016)) are included. The data allows a similar distinction for the content of knowledge as for HEIs into computer science, STEM, social science and economics as well as medicine and interdisciplinary research. The data for both research institutes and HEIs has been aggregated on neighbourhood level (see 5.4 for a Map).

5.3.4 Location Characteristics

The third part contains neighbourhood characteristics from various data sources. The number of commercial buildings and the number of private households is used to capture the firms' local environment, such as the distinction between suburban office parks and dense neighbourhoods (Breidenbach and Eilers, 2018). Investigating the role of transportation infrastructure, data from OpenStreetMap (OSM) covering the number of motorway accesses, light rail stations and bus stops is used (OpenStreetMap, 2018). Further, the distance to the Cental Business District (CBD) has been calculated for each grid taking the city hall as center.

The OSM data are not time-varying there is little to no variation over time to expect in transportation infrastructure. Furthermore, the number of bars and restaurants has been retrieved for 2018 as an indicator for urban amenities (OpenStreetMap, 2018). Additionally, I am using the average rent price for a 60 m^2 apartment as a proxy for real estate values and a willingness to pay for amenities as provided in the RWI-GEO-RED dataset. The data originate from ImmobilienScout24, the largest real estate portal in Germany, and is provided by the RWI (RWI and ImmobilienScout24, 2021).

5.3.5 Descriptive Statistics

Table 5.2 provides summary statistics for digital firms and knowledge infrastructure on neighbourhood level for Hamburg, Berlin and Munich covering the years 2009 to 2016. The mean number of new firms per grid is 1.21, while the maximum is 79 indicating a spatial concentration of the firms. Further, there is a maximum of eight research institutes and seven HEI in one grid.[7]

Figure 5.3 show firms' spatial distribution in each city. The graphs on the right hand side show the share of firm birth over time defined as firms founded in the individual grid divided by

[7]Grid-level and neighbourhood-level are used synonymously.

Table 5.2: Summary Statistics Firm-Grid

Statistic	N	Mean	St. Dev.	Min	Max
Firm birth Grid	16,520	1.21	3.75	0	79
Count firms Grid	16,520	6.16	18.71	0	426
Higher Education Institions	16,520	0.04	0.29	0	7
Research Instituions	16,520	0.04	0.31	0	8
Research Universities	16,520	0.01	0.12	0	2
Universities of Applied Sciences (UAS)	16,520	0.03	0.19	0	4
Universities of Design	16,520	0.004	0.08	0	3

Notes: Data covers the years 2009 to 2016. N is the cumulative number of all grids from 2009 to 2016.

all founded firms in the city each year. The neighbourhoods in the maps show the same colors as the graphs, as the five grids with the highest share of firm birth in 2016 are colored in the same order. The Figure depicts a clear tendency of digital firms to over-proportionally cluster in inner city neighbourhoods.

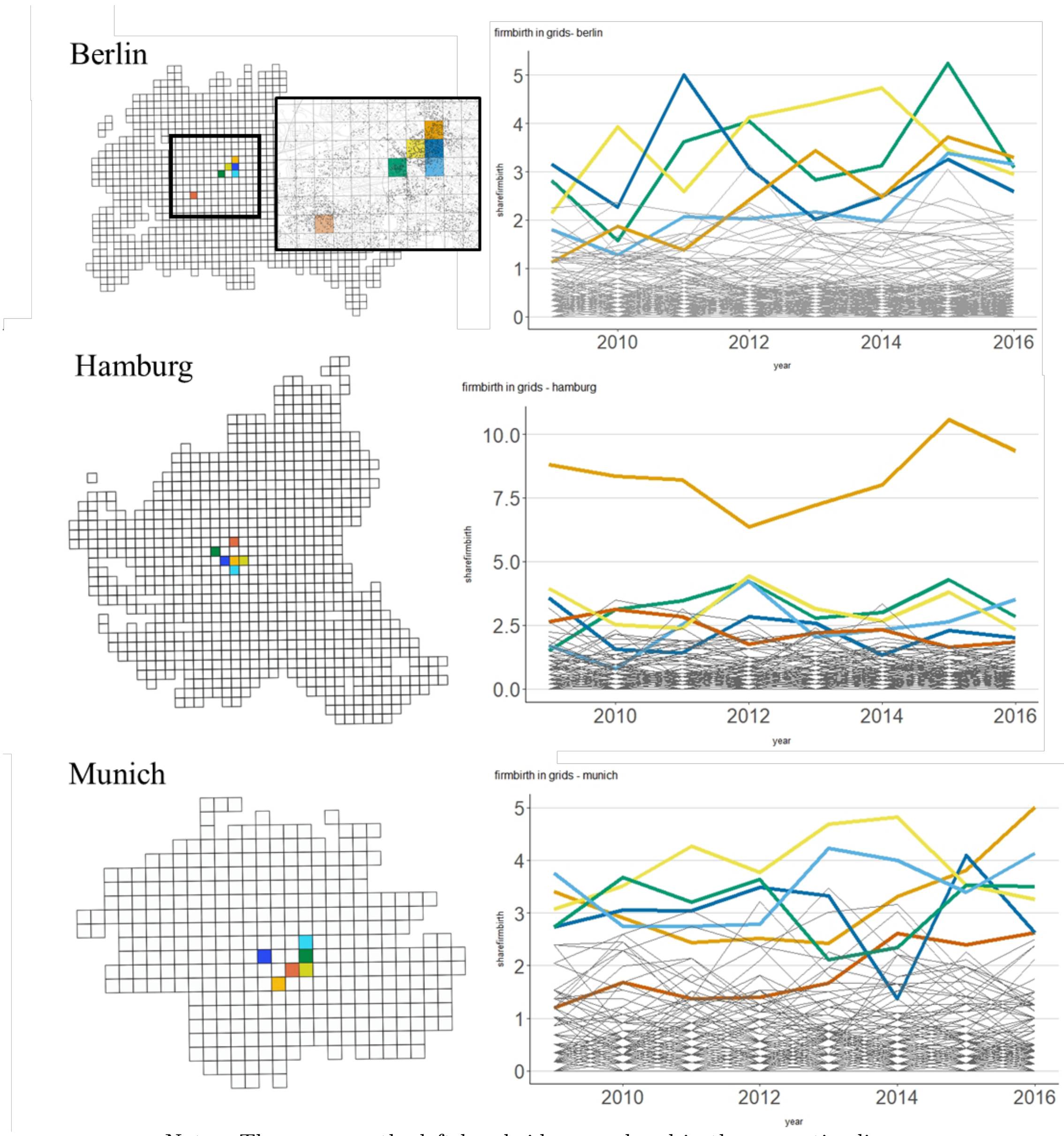
In all cities, four to five grids stand out in firm-density. The top performing grids are all next to each other, except for one spatial outlier in each city. Hamburg seems to be uniquely persistent with a cluster of 10 % of firms in one grid at all times. In Berlin and Munich, the firm birth dynamics vary year by year but only within five grids (ranging from 3-5 % of all firm birth). This could be indicative of distinct location characteristics that are not related to inner-core characteristics like amenities.

Figure 5.4 shows the number of locations for HEIs on the left maps, while the number of research institutes is displayed on the right. In a direct comparison of the top firm birth grids via-a-vis the knowledge institutions, there is an indicative tendency of both to be located in the cities' cores – or at least in close proximity. Welch's unequal variances t-tests show on a 99 % confidence level, that the mean of firm birth in grids with a HEI are significantly different than without a HEI.

5.4 Empirical Strategy

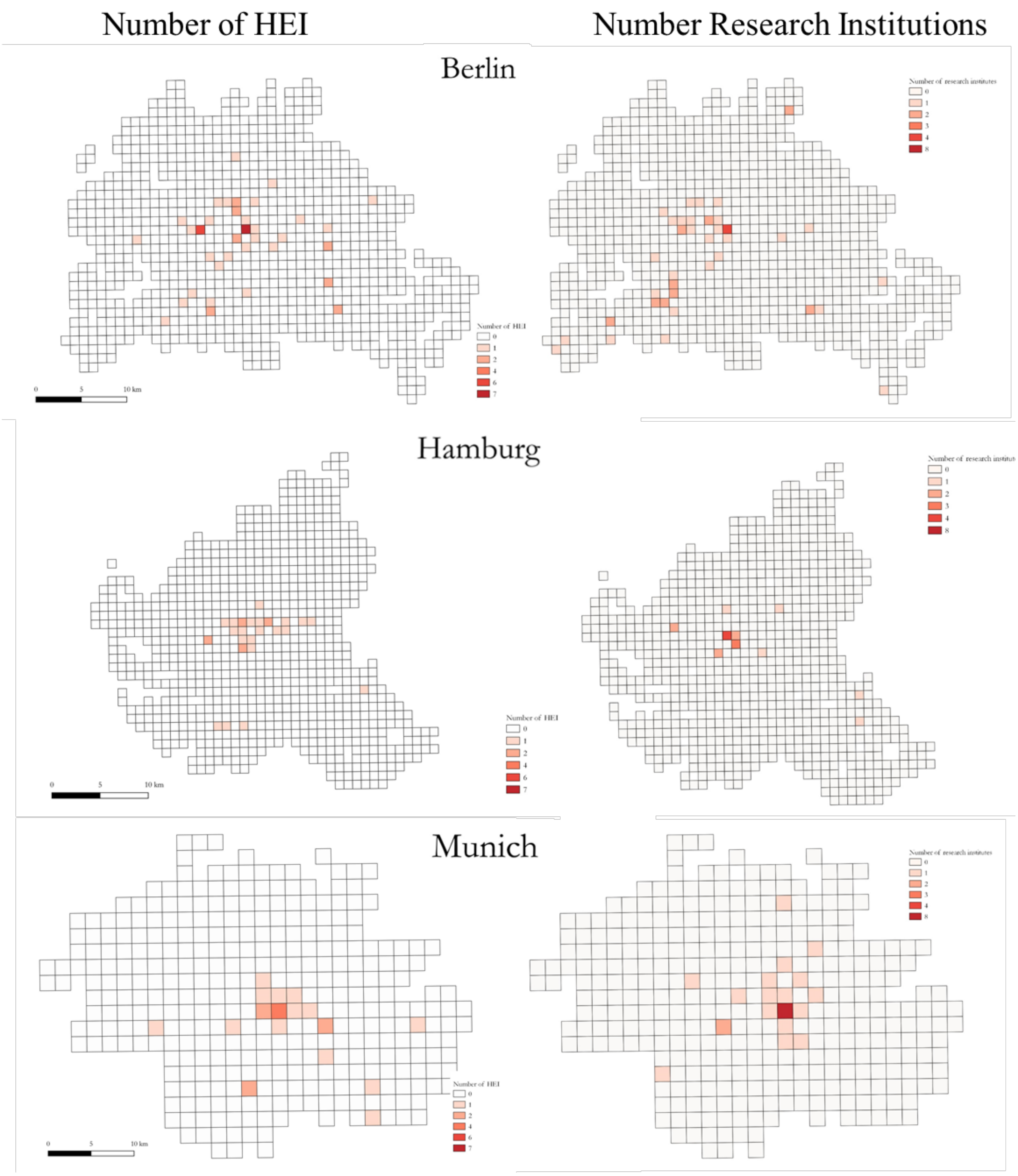
I assume new firms will locate close to one another as well as institutional knowledge sources to capitalize on external knowledge stocks. Therefore, this Chapter uses firm births within neighbourhoods as the dependent variable of the empirical model. The main advantage of using new firms is that those are not constrained by previous location decisions and sunk costs.

Figure 5.3: Digital Firm Birth on Grid-Level for Berlin, Hamburg and Munich



Notes: The maps on the left hand-side are colored in the respective lines of the graphs on the right hand-side. All other grids in grey. Datasource basemap: GeoBasis-D

Figure 5.4: Frequency of Higher Education Institutions in Berlin, Hamburg and Munich



Notes: all other grids in grey. Datasource basemap: GeoBasis-D

Therefore, they provide better information on the role and magnitude of agglomeration effects than existing ones (Gómez-Antonio and Sweeney, 2021). The dependent variable is defined as

$$Y_{n,t} = \Sigma i_{n,t} \tag{5.1}$$

which denotes the sum of all firm births i in neighbourhood n in year t. The uniform 1x1 km squares - the neighbuorhood level - also implies the density per km^2.

The econometric analysis proceeds in three steps: First, I identify general determinants of location of young digital firms with a set of control variables including HEIs and research institutes without further subdivision. Second, I employ a model sensitive towards the institutions' knowledge content, that is, departments and research specialty of the research institutes. Third, I test the HEI's institutional setup in a model distinguishing research universities, UAS and universities of arts, music and design.

For all estimations, I use OLS panel fixed effects models. With the inclusion of time- and city fixed effects, I aim to control for city-specific time-invariant characteristics and yearly increments such as general trends in the industry or the economy at large. The city fixed effects have been chosen over a grid fixed effect to control for individual city characteristics, i.e. a possible 'Berlin Effect'. Additionally, all observable and unobservable effects which might vary on city- and time level are controlled for. This reduces the threat of an omitted variable bias.

The following model will be estimated using OLS with standard errors clustered on neighbourhood level:

$$\begin{aligned}
\ln(Y_{n,t}) = \alpha &+ \ln(STOCK_{n,t-1}) + \ln(STOCK_{n*,t-1}) + \ln(l_{n,t}) \\
&+ HigherEducation_{n,t} + HigherEducation_{n*,t} \\
&+ Research_{n,t} + Research_{n*,t} \\
&+ T_t + \gamma_j + \epsilon_{n,t}
\end{aligned} \tag{5.2}$$

The stock of firms ($STOCK$) and knowledge institutions enter the model on two spatial scales: the neighbourhood level (1x1 km^2 grids n) and the first-order neighbours (3x3 km^2 grids n^*). Values for the first-order neighbours are calculated by summing up the number of observations in the eight first-order neighbours. $l_{n,t}$ refers to the set of locational characteristics in neighbourhood n at time t (commercial buildings, residential buildings, light rail, bus stops, motorway accesses, amenities and prices, distance to CBD, see Section 5.3.4). T_t is a time fixed effect, γ_i is a city fixed effect and $\epsilon_{n,t}$ is the error term. I use the $ln + 1$ for the dependent variable to account for grids without firm birth in a given year. Equation 5.2 provides a baseline bench marked against the literature, as it contains HEIs and research institutes without further subdivision.

To examine the content of knowledge inputs – related or unrelated – being conducive to firm birth, a second model distinguishing departments and contents from research institutes will be estimated

$$\begin{aligned}
\ln(Y_{n,t}) = \alpha &+ \ln(STOCK_{n,t-1}) + \ln(STOCK_{n^*,t-1}) + \ln(l_{n,t}) \\
&+ Departments_{n,t} + Departments_{n^*,t} \\
&+ ResearchContent_{n,t} + ResearchContent_{n^*,t} \\
&+ T_t + \gamma_j + \epsilon_{n,t}
\end{aligned} \tag{5.3}$$

where the sum of departments reflect the number of HEIs in (Equation 5.2) and research institutes respectively. Further, departments of HEIs divide into: IT; STEM; Economics, Social Science and History; Health-/ Medical and Law; and Music, Arts and Design. Research institutes divide into IT; STEM; Economics; Social Science and History; Health/Medical and Interdisciplinary Institutes.

Motivated by the literature on the 'differentiated knowledge base approach' (Trippl et al., 2009) a third model uses the distinction of research universities (analytical knowledge), UAS (synthetic knowledge) and universities for arts, music and design (symbolic knowledge):

$$\begin{aligned}
ln(Y_{n,t}) = \alpha &+ \ln(STOCK_{n,t-1}) + \ln(STOCK_{n^*,t-1}) + \ln(l_{n,t}) \\
&+ ResearchUniversities_{n,t} + ResearchUniversities_{n^*,t} \\
&+ UAS_{n,t} + UAS_{n^*,t} \\
&+ UniversitiesDesign_{n,t} + UniversitiesDesign_{n^*,t} \\
&+ T_t + \gamma_j + \epsilon_{n,t}
\end{aligned} \tag{5.4}$$

5.5 Empirical Results

The following section presents the empirical results. After presenting and discussing the baseline model's results, results of the models concerning knowledge contents –- related vs. unrelated –- are presented and discussed. Third, differences in clustering close to research universities, UAS and universities for music, arts and design are presented.

5.5.1 Baseline model

Table 5.3 reports the coefficients of interest in the baseline model. The coefficients for the location controls are presented in Table A3.2 (Appendix).

Table 5.3: Results of the Baseline Estimation (Model 1)

Baseline Model		
	ln(firmbirth)	
n ln(STOCK_LAG)	0.289***	(0.007)
n Research Inst.	0.029**	(0.014)
n Higher Educaion Inst.	0.104***	(0.014)
n^*ln(STOCK_LAG)	0.036***	(0.004)
n^* Research Inst	0.012**	(0.005)
n^* Higher Educaion Inst.	0.048***	(0.006)
Neighbourhood controls	Yes	
City FE	Yes	
Time FE	Yes	
Observations	12,726	
Adjusted R2	0.700	

Notes: Cluster Robust Standard Errors in parenthesis on grid-level. Significant levels at ***p < 0.001; **p < 0.01; *p < 0.05

The baseline assumption is that digital firms cluster within cities. The model confirms significant firm clustering as a one percent increase in the stock of firms is followed by a 0.289 percentage point increase in firm births. This indicates the existence of localisation economies operating in neighbourhoods of one square kilometer and less which is in line with Arzaghi and Henderson (2008), Andersson et al. (2019) as well as Lavoratori and Castellani (2021). This shows that localised within-industry spillovers are mainly tacit with a great importance of face-to-face contacts. This effect is underpinned by the finding that clusters decay sharply with distance as the effect for the stock of firms in n^* is only 12 % of the effect in the direct neighbourhood n. On this basis, the inner-city location of the clusters can now be used to assess the relevance of other location factors. If HEI-industry links are as tacit as within-industry spillovers, clusters of digital firms occur in close proximity to HEIs and research institutes within cities (Hypothesis 1).

Results reveal significant firm clustering close to HEIs. One additional HEI on average leads to a 10.4 % increase of digital firm birth within the neighbourhood (n) when controlling for other location characteristics. Similarly, for an additional research institute, digital firm birth rises by 2.9 %. In contrast to Duvivier et al. (2018), institutional knowledge does not seem to impact the city's parameter homogeneously as in line with Rammer et al. (2020).

However, these results go beyond digital firm clustering close to external knowledge sources. When assessing a relative importance, the effect for an additional HEI and almost three times larger than for research institutes. This is probably due to disproportionately large political focus and funding for universities over research institutes. Universities' 'third mission' besides research and teaching provides knowledge and technology transfers in economically useful ways to contribute to (local) economic growth and prosperity (Bercovitz and Feldman, 2006).

However, the effects for institutional knowledge do not decay as sharply with distance as within-industry. For research institutes, the effect is still at 41 %, while being at 46 % for HEIs in n^*. This indicates that advantages from closeness to knowledge institutes do not diminish as rapidly with distance as for industry knowledge. Thus, university-industry channels are less sensitive to proximity than intra-industry channels. The disadvantage of not having a face-to-face contact is smaller for university-industry dynamics than within the industry. Thereby, more firms can possibly gain advantages from HEIs knowledge independent locations within the city. This, in turn, partly supports Duvivier et al. (2018), who suggest that HEI's impact goes beyond the neighbourhood levels.

5.5.2 Related and unrelated knowledge contents

Model 2 investigates what tacit knowledge contents digital firms' gain from HEIs and research institutes vis-a-vis the localisation economies gained from similar firms nearby. Results on the question whether digital firms cluster in neighbourhoods where related knowledge can be retrieved from HEIs and research institutes (H2) are discussed. Table 5.4 presents an overview of the results, Table A3.3 (see Appendix) contains the full regression results.

Table 5.4: Regression Results: Knowledge Contents

		Firm birth in n	Firm birth in n*
Related knowledge	IT –Research	/	/
	IT – HEI	0.206***	/
	ECON – Research	0.090***	0.024***
	ECON – HEI	/	0.021***
	Design – HEI	0.167***	0.080***
Unrelated knowledge	Interdisciplinary Research	/	/
	Med – Research	/	0.067***
	Med – HEI	/	0.091***
	STEM Research	/	/
	STEM – HEI	/	-0.093***

Notes: The table only presents the very basic results of one regression: The "/" indicates no significant effect. Significant levels at ***p < 0.001; **p < 0.01; *p < 0.05. For the full regression table see Appendix Table A3.3, adjusted r^2 is 0.703

Departments in HEIs and research institutes for IT and data science; departments and research institutes for businesses and economics (the measurement here also includes other social, political or historical science departments) and design schools are defined as institutional knowledge sources of related variety for digital firms (Pina and Tether, 2016). Results indeed show significant firm clustering in neighbourhoods where such related knowledge is institutionally bounded. There is a significant effect of HEI for IT and data science and of research institutes for social sciences and for design departments.

However, the effects vary significantly in size: An additional data science department is followed by an average of 20 % more firm birth, while the effect shrinks to 16 % for a design school and to only 9 % for a research institute on social sciences. This means that localised knowledge in the sense of the agglomeration literature has the strongest effect on firm clustering, whereas directly compared, the effect of related knowledge in the narrower sense is weaker, but still remarkably strong. Further, this result is in line with Trippl et al. (2009) as it indicates that different types of knowledge do not need to be retrieved in the same quantity. Overall, the results indicate that for each of these knowledge contents, tacit knowledge transfers are advantageous for digital firms.

The finding that digital firms cluster in neighbourhoods where related knowledge can be retrieved from HEIs and research institutes is additionally underpinned by the fact that there is no significant firm clustering for unrelated knowledge contents on neighbourhood level that is STEM, Medicine and Law, and interdisciplinary research.

However, including the first-order neighbours (n^*), results are similar to the findings on the spatial decay within the industry, that the effect for an IT and data-science HEI decays that sharply that it is insignificant for n^*. Similarly, the effect for design and creative schools is less than 50 % compared to the direct neighbourhood. Conversely, the clustering effect turns positive and significant for HEIs on social sciences for n^* as well as for highly specific medical and legal HEIs and research institutes. Taken together, these findings indicate that economically proximate knowledge requires spatial proximity. For less economically proximate knowledge, transfer channels other than tacit exchange suffice. This, however, has not been shown on such small spatial scales in the literature.

5.5.3 Types of HEI - Differentiated knowledge bases

The previously presented results indicate that the importance of spatial proximity for knowledge transmission can vary given the economic proximity and similarity of knowledge from potential knowledge exchanging subjects. In other words, the relevance of tacit knowledge and face-to-face contacts is determined by the knowledge that is to be exchanged. Based on the differentiated knowledge base approach, firms are expected to locate closer to symbolic (design universities) than analytical (research universities) and synthetic (UAS) knowledge bases (H3). Table 5.5 presents the results.

Table 5.5: Regression Results Knowledge Bases

Knowledge Bases in Types of Higher Education Institutions		
	ln(firmbirth)	
n Research Universities	-0.029	(0.044)
n Univ. of Applied Sciences	0.140***	(0 .024)
n Univ. Arts/Music/Design	0.165***	(0.046)
n* Research Universities	0.036**	(0.016)
n* Univ. of Applied Sciences	0.049***	(0.008)
n* Univ. Arts/Music/Design	0.066***	(0.024)
Neighbourhood Controls	Yes	
City FE	Yes	
Time FE	Yes	
Observations	12,726	
Adjusted R2	0.701	

Notes: Cluster Robust Standard Errors in parenthesis on grid-level. Significant levels at ***p < 0.001; **p < 0.01; *p < 0.05

Results show significant firm clustering for UAS and design schools on neighbourhood level but no significant effect for research universities. Considered by itself, this is a first indication that analytical knowledge can be transferred over longer distances while the relevance of face-to-face contacts is higher for symbolic and synthetic knowledge. This implicitly supports that tacit knowledge is of higher value for synthetic (an average of 16.5 % increase in firm birth for an additional HEI) than for symbolic knowledge (14 % increase).

Further to this, the positive effect of synthetic and symbolic knowledge decays with distance, while the effect for analytical knowledge increases with distance (see results for n^*). Still, the effect for digital firm birth for symbolic knowledge is almost double the effect of analytical knowledge. Nevertheless, the effect for synthetic engineering knowledge decays sharper (by 65 %) relative to symbolic knowledge (60 %).

The results provide partial evidence of Hypothesis 3 as differences in the spatial sensitivity to the knowledge bases clearly exist. In line with e.g Asheim and Gertler (2005), Moodysson (2008) and Plum and Hassink (2011); analytical knowledge is the least distance-sensitive knowledge base. Regarding synthetic and symbolic knowledge, the effects are not as clear-cut within cities. However, results do not contradict the findings of Trippl et al. (2009) that ICT firms have strong connections to local HEIs with applied research, with prototyping and testing being the most important areas of cooperation.

Additionally, Toner (2010), among others, finds UAS to be especially likely to cooperate with the local industry (compared to research universities). Finally, highly context-specific symbolic knowledge does not travel easily and therefore the value of a face-to-face contact is most important. Hence, this Chapter is the first to empirically test the differentiated knowledge base theory with its spatial reach on micro levels with the finding that spatial transferability of knowledge bases seem to function on much smaller scales than investigated to date.

5.5.4 Robustness

To assess the robustness of the results, several tests have been conducted. For robustness of the cluster-density measure, I employ an alternative measure: the share of firm birth in the neighbourhood (in year t) relative to all firm birth in city y in year t. Additionally, I test relocation patterns by employing the same measures to in-moving firms instead of firm births.

Results for all estimations are presented in Tables A3.2, A3.3 and A3.4 (see Appendix). Results show that firm birth and relocation patterns are remarkably similar. Young digital firms show a continuous need for knowledge externalities as well as external knowledge updates and do not seem to need knowledge from HEIs only at the get-go. In sum, local firm clusters seem to be self-reinforcing as they grow by firm births and in-movers. A concern is the high number of grids that did not register firm birth in a given year (69 % of the grids). The fifth columns in Tables A3.2, A3.3 and A3.4 therefore show the regressions with a subsample of the grids that registered positive firm birth. Results are robust to the baseline with the full sample.

Third, the baseline regression is estimated without fixed effects (see Appendix Table A3.5) for comparison reasons of within and between effects. Results are consistent with the fixed effects model. Fourth, the effect of larger distances to research universities (analytical knowledge) could be biased by big established research universities located on campuses that take on a lot of space within a neighbourhood, leaving little space for firms (as addressed in Duvivier et al. (2018)). The models have been estimated including a dummy variable for campuses (Appendix Table A3.6).[8] Results show that neither the campus dummy nor its first–order neighbour is significant; results in Model 3 are thus robust against this concern.

Further, the underlying definition of digital firms in this Chapter is different to other papers by moving away from standard industry classifications by including digital firms in other industries (for the reasons elaborated above). To ensure the efficiency of the sample, Table A3.7 presents the baseline regression results with a sub-sample of the data containing only firms that are registered in the ICT sector. Results are robust, with the only difference that research institutes are not significant in n for the narrow definition of ICT firms. This further underpins the necessity to loosen up hard lines in separating businesses by their industry codes instead of the knowledge their competitiveness draws on.

5.6 Conclusions

This study shows that neighbourhood-level spatial proximity to knowledge institutions is a significant explanatory factor for firm birth. Specific knowledge that can be transferred into the firm without major detours, that is data-science, business-knowledge and design-knowledge for marketability, contributes to clustering of digital firms within cities.

The contribution to the literature is threefold. First, the Chapter adds to the wide academic debate on localisation vs. urbanisation on micro-scales. The effects from specialisation found in this Chapter might have been detected as a diversity effect in other papers, for example in Andersson et al. (2019), as I employ the underlying knowledge instead of strict industry codes for firm classification. This Chapter shows that knowledge-diversity in a city is not randomly distributed, but location decisions of firms rely on a related variety of knowledge. This gives indication for a localized input-output system that spans firms in different sectors or knowledge institutions that share the same ultimate knowledge. Therefore, future work should deepen the

[8]Google Maps has been used to retrieve whether the HEI is organised as a campus or in a single building. A location is considered a campus if it consists of more than three buildings and hosts more than 30 degree programs.

analysis on fine shades of specialisation and diversity to gain deeper understandings for urban input-output systems.

Second, the Chapter links industry dynamics and knowledge institutions because the latter are one of the main vehicles in political strategies to foster regional growth. In the face of public money invested in knowledge transfers, it is important to investigate whether firms seem to gain advantages from close spatial proximity. If this is the case, public funds would disproportionally hit a selected group of firms competitive enough to secure office space in the right radius around a knowledge institution. Results show that research universities do not contribute to firm clustering in their direct neighbourhoods, while UAS and universities for music, arts and design do. This indicates transfer channels from research universities are more likely to affect many firms in the region, while effects from applied and creative universities act more locally. Therefore, the mere geographical presence of knowledge in space is probably not sufficient for knowledge exchanges.

The third contribution to the state of the art literature is the differentiated knowledge base approach. As until now, authors mainly investigated (inter)national, regional and local levels, this Chapter shows that the transferability of the knowledge bases seem to function on much more granular scales than investigated until now.

Overall, the Chapter shows when taking the city as a molecule in the economic ecosystem, it seems that knowledge serves as the electromagnetic force that determinants the position of atoms – a complex system of firms and institutions.

Chapter 6

Summary and Concluding remarks

6.1 Summary of the results

The aim of the thesis was to provide a detailed analysis of the evolution of the digital industry in Germany with regard to the spatial dimension. This topic is of particular interest for policy makers on several spatial levels. Digital firms fuel economic growth and contribute to the transition to a knowledge based economy along the lines of innovation and entrepreneurship to ensure future competitiveness. As digital services complement to almost all other sectors, the political interest of attracting those firms is equally relevant for policy makers on national, regional and local scales.

To foster entrepreneurship-based growth of the digital industry and thus provide tailored policies, it is key to understand the underlying economic mechanisms and factors that contribute to a vital firm-environment. Therefore, the evolutionary processes of the German digital industry are considered on several spatial levels in this dissertation, putting the development into perspective. The results yield insights into agglomeration effects and spatially bounded knowledge flows captured by entrepreneurial firms beyond the mere description of the industry development.

Building on the theoretical backbone of the geography of innovation, resulting in entrepreneurship and economic growth such as agglomeration economies, the first empirical contribution in Chapter 3 begins with investigating location patterns of firm birth and relocation on NUTS 3 level. This is the broadest spatial scale applied in this dissertation and mostly benchmarks core-cities against their urban peripheries and more rural areas, as it covers administrative independent cities and the administrative merger of smaller cities and municipalities into counties.

The third Chapter measures digital firm birth intensity using a linear regression model (OLS) with county- and time- fixed effects and a host of regional control variables. Results show that accessibility of industry-specific knowledge as proxied by the co-location of digital firms is highly conducive to startup activity. In line with the literature, universities are a significant factor for new digital businesses. This result indicates that infant digital firms rely on, and potentially originate from industry as well as institutional know-how and that locational costs outweigh such locally-bound, tacit benefits as laid out in the Introduction (see Chapter 1.1.2).

Further, the contribution presented in Chapter 3 reveals insights into the relocation of young firms, that is spatial shifts in the industry as a whole. The contribution to the literature lies in the usage of a fixed-effects gravity model using aggregate relocation flows while simultaneously considering the spatial dimension captured by the distance.

The results first show that significantly more digital businesses relocate to counties with a high density of digital firms. Therefore, digital firms show a strong preference to cluster. Further, flows between neighbouring counties are more than twice as big as other flows. In other words, counties receive more than twice as many relocates from their neighbouring counties than from others. Additionally, relocation flows decay with distance. This is consistent with the fact that moving costs as captured by distance play a crucial, deterring role in relocation decisions. It indicates that digital firms tend to stay in their regions of origin. Accordingly, firms with high dependency on outside resources and strong networks as it should be the case for young firms in the digital industry do not relocate over long distances.

The dominance of short-distance relocations highlights the strong regional persistence of entrepreneurship in general and the digital industry in particular. Core-cities offer breeding grounds for new businesses and the surrounding counties benefit once these firms decide to leave their birth towns. On a more abstract level, the results imply that firms are willing to access networks and knowledge that is bound in the region after relocation and thereby manifest entrepreneurship capital.

For regional policy makers, that shows that targeting digital firm birth will also spill over into neighbouring counties in the medium to long run next to the expected initial local benefits. Thus, competitiveness-improving effects are not limited to a certain region but spread across NUTS 3 borders to a limited extend. Therefore, an intra-regional cooperation strategy to foster firm birth where counties administrations pool their resources would be a promising approach.

Motivated by these results, Chapter 4 analyses patterns within coherent labour market regions to further investigate location factors of the digital industry in core-periphery dynamics.

As the results of Chapter 3 clearly show that core-cities are key drivers of the development, Chapter 4 therefore focuses on the understudied municipalities surrounding core-cities. It is in the interest of municipalities' decision makers to attract the entrepreneurship capital of their regions not only as an attractive location for moving firms, but for firm birth as well. The chapter shows that the main advantage of a municipality within an urban labour market region is essentially its location, that is the distance to the next core city. This empirically underpins the important role of agglomeration effects and sector-specific factors such as networks and the accessibility of geographically bound knowledge in the theoretical regional models. When comparing results from NUTS 3 vis-à-vis LAU regions, it becomes evident that the county level only allows a very broad picture, as they absorb smaller, second-tier cities. Further, the spatial decay of externalities deriving from big core-cities are not equally distributed in counties, but municipalities being located close to the cities host more advantages. Besides these contributions, Chapter 4 provides an empirical contribution to the literature on mono-and polycentric urban areas, and sheds light on whether agglomeration externalities diffuse differently in space when there is more than one core that spills over (dis-)advantages. This is equally relevant for local and regional policy makers, as a deep understanding of competition effects of a knowledge intensive industry may request re-pivoting of regional specialisation profiles.

Results show that urban cores in monocentric urban regions tend to absorb entrepreneurship capital from their direct neighbours with growing industries, superimposing an 'agglomeration shadow' over their direct neighbours. Yet, municipalities in monocentric urban regions gain advantages from population growth and universities of the cores. Successful cores in polycentric urban regions however serve their neighbours by increasing their relative attractiveness over other municipalities and allow them to borrow industry-specific externalities. This gain in polycentric urban regions is specific to the industry, as population growth and institutionalised knowledge in additional universities are not significant for polycentric urban regions. On top of that, this indicates competition effects of the individual centers within polycentric urban regions.

The results of Chapter 3 and Chapter 4 highlight the role of knowledge and the accessibility of knowledge for firm birth and in later stages of the digital firms' life cycle. This is in line with the theoretical considerations as laid out in the Introduction on the geographical dimension of knowledge and its links to innovation and entrepreneurship (see Chapter 1.1.2). The digital firms show a clear tendency to cluster regionally in and around big cities that provide a host of externalities and amenities. Motivated by this spatial sorting into cities, Chapter 5 investigates digital firms location decisions within cities. Due to the identification of knowledge, transferred by within industry tacit spillovers from similar firms, as well as the general role of universities and research institutes as knowledge institutions, the empirical analysis links within city clustering of firms and proximity to knowledge institutions.

The aim of this final contribution is to shed further light on possible transmission channels of knowledge flows into firms. The results show indeed a clustering of digital firm birth on neighbourhood-level in close spatial proximity to knowledge institutions.

Nonetheless, significant firm birth patterns do not occur close to random neighbourhoods that host any kind of Higher Education Institutions, but in those where specific industry relevant knowledge is present. That is knowledge that can be transferred into the firm without major detours: data-science, business-knowledge and design-knowledge for marketability. More diverse knowledge inputs are only significant for first-order neighbourhoods, defined as a 3x3 km^2 scale. This contributes to the economic literature by showing that advantages from localisation (same industry spillover) work on smaller scales than advantages from urbanisation (diverse spillover, e.g. Andersson et al. (2019)).

Moreover, the insights of Chapter 5 yield results regarding whether the institutional settings of the knowledge stored in knowledge institutions have an effect on micro-level digital firm birth. The novelty of this approach lies in linking the empirical setup to the literature on the 'differentiated knowledge base approach', which has been predominantly used in the planning literature and tested in qualitative research designs only. The intellectual baseline is that highly codified analytical knowledge (as stored in research universities) can be transferred over long distances. Universities of Applied Science rather host more engineering-based, problem solving, applied synthetic knowledge that is not transferred over long distances as easily as analytical knowledge. Finally, symbolic, highly context specific knowledge requires face-to-face contacts when transmitted and is mostly found in universities of arts. I find significant firm clustering in neighbourhoods with Higher Education Institutions and research institutes. However, cluster effects within the industry decay more rapidly over distance (88 %, which means the effect of the number of firms within the same neighbourhood is 88 % higher than for contiguous neighbourhoods) than for Higher Education Institutions (54 %) and research institutes (59 %). This indicates tacit spillovers from knowledge institutions to be conducive to new firm births. At the same time, transmission channels less sensitive to geography (e.g. networks), seem to be stronger than for within-industry linkages. Of course, these are only proxies to measuring knowledge flows, because especially tacit spillovers are hard to capture, particularly over long time spans.

After all, the commercialisation of new knowledge is revealed in founding a business in a specific location.

Overall, Chapter 5 shows a significant clustering of firm birth activity within cities. This result on its own is of great importance for local policy makers. Providing inner city locations that meet preferences of entrepreneurs while ensuring a vivid knowledge exchange between them is important for the local industry growth. This is mainly a task for urban development and business promotion units. However, subsidizing office spaces or financing start up hubs in the immediate vicinity to research universities is, at least for the digital industry, less fruitful than considering locations next to more tacit knowledge sources.

The finding that research universities in particular seem to channel its knowledge on broader spatial scales comes with advantages. In line with the findings of Chapter 3 and 4, the contribution of the knowledge stored in the universities is not limited to very few neighbourhoods, but contributes to a general positive development of the city and its neighbours.

6.2 Concluding remarks

When jointly considering the three individual empirical contributions presented in this dissertation, the main takeaway is that the evolution of the digital industry in Germany is clearly an urban phenomenon with a strong regional persistence that is shaped by the geography of knowledge.

In light of the theoretical foundations of the 'Knowledge Spillover Theory of Entrepreneurship' (Acs et al., 2009), entrepreneurs commercialise on new ideas originating in incumbent firms or Higher Education Institutions and research institutes. The results of all three empirical contributions show that co-location of similar firms drawing on the same knowledge base as captured in the core-dataset and the presence of knowledge institutions are determinants for additional firm birth on all three investigated spatial levels. One major caveat of the analysis of location choices of the firms is, that the channels of knowledge transfer cannot be disentangled in more detail. Thus, it is not entirely clear that the location choice of the new firm is indeed close to the incumbent firm. Analyzing the location choice reveals the preference of a profit-maximising firm that is reliant on outside resources. While there is no claim for causality, results nevertheless indicate that the chosen locations are advantageous for the firms. The inclusion of location decisions of maturing firms by tracing firm relocations contributes to show that firms desire these knowledge inputs in the long run.

The results of all contributions partly contribute to the scholarly debate on whether firms derive advantages from localisation - same industry externalities - or urbanisation, that is economically diverse environments. The inclusion of firms drawing on a similar business model and thereby knowledge base in the digital sector gives support that Porters' notion of clusters (Porter, 1990) is conducive for a vital breeding ground for new ideas which results in the birth of new firms.

This dissertation leaves some outlooks for future research. First, it would be interesting to introduce measures on the success of new firms linked to its location. Nevertheless, especially for new and young firms, this provides some obstacles. A longer survival of firms indicates a successful business model. This measure requires the consideration of very long time spans, where surviving firms are benchmarked against firms exiting the market. Further, measuring the financial position of firms and the number of employees as a growth indicator are hard to retrieve, as small and new companies are not obliged to report specific measures to the Federal Gazette. Additionally, Bijedić et al. (2020) state that firms with groundbreaking incremental innovations often take longer to become profitable.

Moreover, in the context of the spatial development of regions and cities, it is interesting to investigate political interventions and external funding. This requires rich datasets on (public) funding and precise tracking of investment flows traced down to the firm level. Lastly, future research should also tackle the impacts of strong digital firm clusters on the individual locations. This is particularly interesting for developments within cities. This dissertation shows in how far the digital industry is shaped by local and regional circumstances. A prospect for regional researchers would be to investigate in how far the digital industry (re-)shapes inner city dynamics in the light of today's societal and ecological obstacles that must be overcome in the future.

Bibliography

Acs, Z. J., D. B. Audretsch, and E. E. Lehmann (2013): "The knowledge spillover theory of entrepreneurship," Small Business Economics, 41, 757–774.

Acs, Z. J., P. Braunerhjelm, D. B. Audretsch, and B. Carlsson (2009): "The knowledge spillover theory of entrepreneurship," Small Business Economics, 32, 15–30.

Acs, Z. J. and A. Varga (2005): "Entrepreneurship, agglomeration and technological change," Small Business Economics, 24, 323–334.

Anderson, J. E. (1979): "A theoretical foundation for the gravity equation," The American Economic Review, 69, 106–116.

Anderson, J. E. and E. Van Wincoop (2003): "Gravity with gravitas: A solution to the border puzzle," American Economic Review, 93, 170–192.

Andersson, M., J. P. Larsson, and J. Wernberg (2019): "The economic microgeography of diversity and specialization externalities–firm-level evidence from Swedish cities," Research Policy, 48, 1385–1398.

Anselin, L., A. Varga, and Z. Acs (1997): "Local geographic spillovers between university research and high technology innovations," Journal of Urban Economics, 42, 422–448.

Arauzo Carod, J. M. (2005): "Determinants of industrial location: An application for Catalan municipalities," Papers in Regional Science, 84, 105–120.

Arauzo-Carod, J.-M., D. Liviano-Solis, and M. Manjón-Antolín (2010): "Empirical studies in industrial location: An assessment of their methods and results," Journal of Regional Science, 50, 685–711.

Armington, C. and Z. J. Acs (2002): "The determinants of regional variation in new firm formation," Regional Studies, 36, 33–45.

Arzaghi, M. and J. V. Henderson (2008): "Networking off madison avenue," The Review of Economic Studies, 75, 1011–1038.

Asheim, B. and H. K. Hansen (2009): "Knowledge bases, talents, and contexts: On the usefulness of the creative class approach in Sweden," Economic Geography, 85, 425–442.

Asheim, B. T., R. Boschma, and P. Cooke (2011): "Constructing regional advantage: Platform policies based on related variety and differentiated knowledge bases," Regional Studies, 45, 893–904.

Asheim, B. T. and M. S. Gertler (2005): "The geography of innovation: regional innovation systems," in The Oxford Handbook of Innovation.

Audretsch, D. and M. Keilbach (2004): "Entrepreneurship capital and economic performance," Regional Studies, 38, 949–959.

Audretsch, D. B., W. Bönte, and M. Keilbach (2008): "Entrepreneurship capital and its impact on knowledge diffusion and economic performance," Journal of Business Venturing, 23, 687–698.

Audretsch, D. B., O. Falck, M. P. Feldman, and S. Heblich (2012): "Local entrepreneurship in context," Regional Studies, 46, 379–389.

Audretsch, D. B., D. Heger, and T. Veith (2015): "Infrastructure and entrepreneurship," Small Business Economics, 44, 219–230.

Audretsch, D. B. and M. Keilbach (2007): "The theory of knowledge spillover entrepreneurship," Journal of Management Studies, 44, 1242–1254.

Audretsch, D. B., M. C. Keilbach, E. E. Lehmann, et al. (2006): Entrepreneurship and economic growth, Oxford University Press.

Bade, F.-J. and E. A. Nerlinger (2000): "The spatial distribution of new technology-based firms: Empirical results for West-Germany," Papers in Regional Science, 79, 155–176.

Baumol, W. J. (2004): "Entrepreneurial enterprises, large established firms and other components of the free-market growth machine," Small Business Economics, 23, 9–21.

Bayer, F., B. Pauly, and K. Wohlrabe (2022): "Branchen im Fokus: ITK-Branche," ifo Schnelldienst, 75, 56–59.

BBSR (2020): “Bundesinstitut für Bau-, Stadt- und Raumforschung. Indikatoren und Karten zur Raum- und Stadtentwicklung. INKAR. Bonn,” "https://www.inkar.de/".

BBSR (2022a): “Bundesinstitut für Bau-, Stadt- und Raumforschung. Indikatoren und Karten zur Raum- und Stadtentwicklung. Bevölkerung gesamt,” "https://www.inkar.de/", [accessed on May 27 2022].

——— (2022b): “Laufende Stadtbeobachtung - Raumabgrenzungen,” "https://www.bbsr.bund.de/BBSR/DE/forschung/raumbeobachtung/Raumabgrenzungen/deutschland/regionen/Grossstadtregionen/Grossstadtregionen.html", [accessed on May 27 2022].

Bercovitz, J. and M. Feldman (2006): “Entpreprenerial universities and technology transfer: A conceptual framework for understanding knowledge-based economic development,” The Journal of Technology Transfer, 31, 175–188.

Berger, T. and C. B. Frey (2016): “Did the Computer Revolution shift the fortunes of US cities? Technology shocks and the geography of new jobs,” Regional Science and Urban Economics, 57, 38–45.

Bersch, J., S. Gottschalk, B. Müller, and M. Niefert (2014): “The Mannheim Enterprise Panel (MUP) and Firm Statistics for Germany,” ZEW-Centre for European Economic Research Discussion Paper.

Bertoli, S. and J. F.-H. Moraga (2013): “Multilateral resistance to migration,” Journal of Development Economics, 102, 79–100.

Bertschek, I., T. Niebel, C. Rammer, and M. Seifried (2020): “IKT-Branchenbild 2020: Volkswirtschaftliche Kennzahlen, Innovations-und Gründungsgeschehen,” .

Bijedić, T., J. Löher, S. Nielen, N. Schlömer-Laufen, A. Herrmann, and M. Zimmermann (2020): “Von der Idee bis zum Gewinn-eine empirische Analyse der Entwicklungsprozesse von Neugründungen,” Tech. rep., IfM-Materialien.

Bitkom (2022): “Anzahl der Erwerbstätigen in der ITK- und CE-Branche in Deutschland von 1998 bis 2022 (in 1.000),” Bitkom; Bundesagentur für Arbeit; Bundesnetzagentur; Statistisches Bundesamt, bitkom.org.

Blum, B. S., S. Claro, I. Horstmann, and T. Tombe (2020): "The DNA of New Exporters: Spin-offs and FDI at the Extensive Margin of Trade," American Economic Review: Insights, 2, 397–408.

Bodenmann, B. R. and K. W. Axhausen (2012): "Destination choice for relocating firms: A discrete choice model for the St. Gallen region, Switzerland," Papers in Regional Science, 91, 319–341.

Boschma, R. (2005): "Proximity and innovation: a critical assessment," Regional Studies, 39, 61–74.

Breidenbach, P. and L. Eilers (2018): "RWI-GEO-GRID: Socio-economic data on grid level," Jahrbücher für Nationalökonomie und Statistik, 238, 609–616.

Brouwer, A. E., I. Mariotti, and J. N. Van Ommeren (2004): "The firm relocation decision: An empirical investigation," The Annals of Regional Science, 38, 335–347.

Buczkowska, S. and M. de Lapparent (2014): "Location choices of newly created establishments: Spatial patterns at the aggregate level," Regional Science and Urban Economics, 48, 68–81.

Bundesagentur für Arbeit (2019): "IT-Fachleute," Berichte: Blickpunkt Arbeitsmarkt, 4, 2019.

Bundesagentur für Arbeit (2021): "Beschäftigung nach Wirtschaftszweigen (WZ 2008) hochgerechnete Werte (Monatszahlen) Deutschland, Ländern und Regionaldirektionen (Arbeitsort). Berichtsmonat 30.09.2021," .

Burchardi, K. B., T. Chaney, and T. A. Hassan (2019): "Migrants, ancestors, and foreign investments," The Review of Economic Studies, 86, 1448–1486.

Camagni, R., R. Capello, and A. Caragliu (2015): "The rise of second-rank cities: what role for agglomeration economies?" European Planning Studies, 23, 1069–1089.

Carlino, G. and W. R. Kerr (2015): "Agglomeration and innovation," Handbook of Regional and Urban Economics, 5, 349–404.

Chaney, T. (2008): "Distorted gravity: the intensive and extensive margins of international trade," American Economic Review, 98, 1707–21.

Chesbrough, H. W. (2003): Open innovation: The new imperative for creating and profiting from technology, Harvard Business Press.

City of Munich (2022): "IKT - Informations- und Kommunikationstechnologie." "https://stadt.muenchen.de/infos/medien-informations-kommunikationsbranche-iuk.html", [accessed on March 23 2022].

Combes, P.-P., G. Duranton, and L. Gobillon (2019): "The costs of agglomeration: House and land prices in French cities," The Review of Economic Studies, 86, 1556–1589.

Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012): "The productivity advantages of large cities: Distinguishing agglomeration from firm selection," Econometrica, 80, 2543–2594.

Conroy, T., S. Deller, and A. Tsvetkova (2016): "Regional business climate and interstate manufacturing relocation decisions," Regional Science and Urban Economics, 60, 155–168.

Cornett, A. P. (2009): "Aims and strategies in regional innovation and growth policy: A Danish perspective," Entrepreneurship and Regional Development, 21, 399–420.

Delgado, M., M. E. Porter, and S. Stern (2010): "Clusters and entrepreneurship," Journal of Economic Geography, 10, 495–518.

——— (2014): "Clusters, convergence, and economic performance," Research Policy, 43, 1785–1799.

Destatis (2021): "Small and medium-sized enterprises." https://www.destatis.de/EN/Themes/Economic-Sectors-Enterprises/Enterprises/Small-Sized-Enterprises-Medium-Sized-Enterprises/_node.html;jsessionid=FC99D6F5B8376C36E439306517CBDA44.live742, accessed on October 15 2021.

Destatis (2021): "Umsatzsteuerstatistik (Voranmeldungen). 73311-0011," "https://www.genesis.destatis.de/genesis/online?operation=ergebnistabelleQualitaetSeparatAUS&levelindex=2&levelid=1633370000001&downloadname=73311-0011#abreadcrumb",.

Dijkstra, L., E. Garcilazo, and P. McCann (2013): "The economic performance of European cities and city regions: Myths and realities," European Planning Studies, 21, 334–354.

Drezner, T. and Z. Drezner (2001): "A note on applying the gravity rule to the airline hub problem," Journal of Regional Science, 41, 67–72.

Duranton, G. and D. Puga (2001): "Nursery cities: urban diversity, process innovation, and the life cycle of products," American Economic Review, 91, 1454–1477.

——— (2004): "Micro-foundations of urban agglomeration economies," in Handbook of Regional and Urban Economics, Elsevier, vol. 4, 2063–2117.

Duvivier, C., M. Polèse, and P. Apparicio (2018): "The location of information technology-led new economy jobs in cities: office parks or cool neighbourhoods?" Regional Studies, 52, 756–767.

Egger, P. and M. Pfaffermayr (2004): "Distance, trade and FDI: a Hausman–Taylor SUR approach," Journal of Applied Econometrics, 19, 227–246.

ESPON, E. (2017): "Shaping new policies in specific types of territories in Europe: islands, mountains, sparsely populated and coastal regions Luxemburg: ESPON EGTC." https://www.espon.eu/sites/default/files/attachments/ESPON_policy_brief_territories_with_geographic_specificities.pdf".[accessed 25 July 2022].

Esteve-Pérez, S., F. Pieri, and D. Rodriguez (2018): "Age and productivity as determinants of firm survival over the industry life cycle," Industry and Innovation, 25, 167–198.

European Commission (1999): Committee on Spatial Development: ESDP-European Spatial Development Perspective: Towards Balanced and Sustainable Development of the Territory of the European Union: Agreed at the Informal Council of Ministers Responsible for Spatial Planning in Potsdam, May 1999, Office for Official Publications of the European Communities.

——— (2011): Territorial Agenda of the European Union 2020. Towards Inclusive, Smart and Sustainable Europe of Diverse Regions. Agreed at the Informal Meeting of Ministers responsible fr Spatial Planning and Territorial Development on 19th May 2011 Gödöllo, HungaryCommittee, Office for Official Publications of the European Communities.

European Union (2011): European Commission; Directorate-General for Regional Policy: Cities of tomorrow: Challenges, visions, ways forward, Publications Office of the European Union.

Fabiano, G., A. Marcellusi, and G. Favato (2020): "Channels and processes of knowledge transfer: How does knowledge move between university and industry?" Science and Public Policy, 47, 256–270.

Federal Ministry of Housing, U. D. and Construction (2022): "Regiopolen und Regiopolregionen für Deutschland – Begleitforschung des Bundes." https://www.region-gestalten.bund.de/Region/DE/einblicke/2201/regiopolen-regiopolregionen.html [accessed 26 July 2022].

Feldman, M., J. Francis, and J. Bercovitz (2005): "Creating a cluster while building a firm: Entrepreneurs and the formation of industrial clusters," Regional Studies, 39, 129–141.

Feldman, M. and M. Storper (2018): "Economic growth and economic development: Geographical dimensions, definition, and disparities," The New Oxford Handbook of Economic Geography, 143.

Figueiredo, O., P. Guimaraes, and D. Woodward (2002): "Home-field advantage: location decisions of Portuguese entrepreneurs," Journal of Urban Economics, 52, 341–361.

Foord, J. (2013): "The new boomtown? Creative city to Tech City in east London," Cities, 33, 51–60.

Fossen, F. M. and T. Martin (2018): "Entrepreneurial dynamics over space and time," Regional Science and Urban Economics, 70, 204–214.

Fraunhofer Institut (2019): "Institutes and Research Establishments in Germany," `"https://www.fraunhofer.de/en/institutes/institutes-and-research-establishments-in-germany.html"`, [accessed on May 25 2019].

Frenken, K., F. Van Oort, and T. Verburg (2007): "Related variety, unrelated variety and regional economic growth," Regional Studies, 41, 685–697.

Fritsch, M. (2017): "The theory of economic development–An inquiry into profits, capital, credit, interest, and the business cycle," Regional Studies, 51, 654–655.

Fritsch, M. and R. Aamoucke (2013): "Regional public research, higher education, and innovative start-ups: An empirical investigation," Small Business Economics, 41, 865–885.

——— (2017): "Fields of knowledge in higher education institutions, and innovative start-ups: an empirical investigation," Papers in Regional Science, 96, S1–S27.

Fritsch, M. and P. Mueller (2008): "The effect of new business formation on regional development over time: the case of Germany," Small Business Economics, 30, 15–29.

Fritsch, M. and D. J. Storey (2014): “Entrepreneurship in a regional context: Historical roots, recent developments and future challenges,” Regional Studies, 48, 939–954.

Fritsch, M. and M. Wyrwich (2014): “The long persistence of regional levels of entrepreneurship: Germany, 1925–2005,” Regional Studies, 48, 955–973.

——— (2017): “The effect of entrepreneurship on economic development—an empirical analysis using regional entrepreneurship culture,” Journal of Economic Geography, 17, 157–189.

Geuna, A. and A. Muscio (2009): “The governance of university knowledge transfer: A critical review of the literature,” Minerva, 47, 93–114.

Ghinamo, M. L. (2012): “Explaining the variation in the empirical estimates of academic knowledge spillovers,” Journal of Regional Science, 52, 606–634.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman, and A. Shleifer (1992): “Growth in cities,” Journal of Political Economy, 100, 1126–1152.

Glaeser, E. L., S. P. Kerr, and W. R. Kerr (2015): “Entrepreneurship and urban growth: An empirical assessment with historical mines,” Review of Economics and Statistics, 97, 498–520.

Glaeser, E. L. and W. R. Kerr (2009): “Local industrial conditions and entrepreneurship: how much of the spatial distribution can we explain?” Journal of Economics & Management Strategy, 18, 623–663.

Glaeser, E. L., S. S. Rosenthal, and W. C. Strange (2010): “Urban economics and entrepreneurship,” Journal of Urban Economics, 67, 1–14.

Gómez-Antonio, M. and S. Sweeney (2021): “Testing the role of intra-metropolitan local factors on knowledge-intensive industries’ location choices,” The Annals of Regional Science, 66, 699–728.

Good, M., M. Knockaert, B. Soppe, and M. Wright (2019): “The technology transfer ecosystem in academia. An organizational design perspective,” Technovation, 82, 35–50.

Grillitsch, M., R. Martin, and M. Srholec (2017): “Knowledge base combinations and innovation performance in Swedish regions,” Economic Geography, 93, 458–479.

Grillitsch, M., T. Schubert, and M. Srholec (2019): “Knowledge base combinations and firm growth,” Research Policy, 48, 234–247.

Guzman, J. (2019): “Go west young firm: agglomeration and embeddedness in startup migrations to Silicon Valley,” Columbia Business School Research Paper.

Head, K. and T. Mayer (2014): “Gravity equations: Workhorse, toolkit, and cookbook,” in Handbook of International Economics, Elsevier, vol. 4, 131–195.

Helmholz Gesellschaft (2019): “Our Research Centers,” "https://www.helmholtz.de/en/about-us/helmholtz-centers/", [accessed on May 20 2019].

Helpman, E., M. Melitz, and Y. Rubinstein (2008): “Estimating trade flows: Trading partners and trading volumes,” The Quarterly Journal of Economics, 123, 441–487.

Henderson, J. and S. Weiler (2010): “Entrepreneurs and job growth: probing the boundaries of time and space,” Economic Development Quarterly, 24, 23–32.

High-Tech Gründerfonds (2021): “HTGF Fonds,” "https://www.htgf.de/de/ueberuns/fonds/", [accessed on 18 January 2021].

Hochschulkompass (2020): “Download der Hochschullisten,” "https://www.hochschulkompass.de/hochschulen/downloads.html", [accessed on June 03 2020].

——— (2022): “Hochschultypen,” "https://www.hochschulkompass.de/hochschulen/hochschullandschaft/hochschultypen.html", [accessed on March 14 2022].

Holl, A. (2004): “Start-ups and relocations: Manufacturing plant location in Portugal,” Papers in Regional Science, 83, 649–668.

Jacobs, J. (1969): “The Economics of Cities Random House,” New York, NY.

Jang, S., J. Kim, and M. von Zedtwitz (2017): “The importance of spatial agglomeration in product innovation: A microgeography perspective,” Journal of Business Research, 78, 143–154.

Jones, C. (2002): “Introduction to Economic Growth, New York: WW Norton.” .

Karemera, D., V. I. Oguledo, and B. Davis (2000): “A gravity model analysis of international migration to North America,” Applied Economics, 32, 1745–1755.

Kerr, W. R. and S. D. Kominers (2015): “Agglomerative forces and cluster shapes,” Review of Economics and Statistics, 97, 877–899.

Kim, J., T. Kollmann, A. Palangkaraya, and E. Webster (2022): "Does local technological specialisation, diversity and dynamic competition enhance firm creation?" Research Policy, 51, 104557.

Klepper, S. (2002): "The capabilities of new firms and the evolution of the US automobile industry," Industrial and Corporate Change, 11, 645–666.

Klick, L. and S. Schaffner (2021): "Regional Real Estate Price Indices for Germany, 2008–2019: RWI–GEO–REDX," Jahrbücher für Nationalökonomie und Statistik, 241, 119–129.

Knoben, J. (2011): "The geographic distance of relocation search: an extended resource-based perspective," Economic Geography, 87, 371–392.

Kohler, H.-P. (1997): "The effect of hedonic migration decisions and region-specific amenities on industrial location: Could Silicon Valley be in South Dakota?" Journal of Regional Science, 37, 379–394.

Kollmann, T., P. B. Jung, L. Kleine-Stegemann, J. Ataee, and K. de Cruppe (2020): "Deutscher Startup Monitor 2020: Innovation statt Krise," Frankfurt am Main: PWC.

Krehl, A. and S. Siedentop (2019): "Towards a typology of urban centers and subcenters–evidence from German city regions," Urban Geography, 40, 58–82.

Kronenberg, K. (2013): "Firm relocations in the Netherlands: Why do firms move, and where do they go?" Papers in Regional Science, 92, 691–713.

Krugman, P. (1991): Geography and Trade, MIT Press.

Larsson, J. P. (2014): "The neighborhood or the region? Reassessing the density–wage relationship using geocoded data," The Annals of Regional Science, 52, 367–384.

Lavoratori, K. and D. Castellani (2021): "Too close for comfort? Microgeography of agglomeration economies in the United Kingdom," Journal of Regional Science, 61, 1002–1028.

Lee, I. H. I. (2020): "Startups, relocation, and firm performance: a transaction cost economics perspective," Small Business Economics, 1–20.

Lee, Y. (2008): "Geographic redistribution of US manufacturing and the role of state development policy," Journal of Urban Economics, 64, 436–450.

Leibniz Association (2019): “Leibnitz Institutes,” https://www.leibniz-gemeinschaft.de/en/institutes/leibniz-institutes-all-lists.html", [accessed on May 20 2019].

Leone, R. A. and R. Struyk (1976): “The incubator hypothesis: evidence from five SMSAs,” Urban Studies, 13, 325–331.

Lucas Jr, R. E. (1988): “On the mechanics of economic development,” Journal of Monetary Economics, 22, 3–42.

Mameli, F., A. Faggian, and P. Mccann (2014): “Estimation of local employment growth: Do sectoral aggregation and industry definition matter?” Regional Studies, 48, 1813–1828.

Manjón-Antolín, M. C. and J.-M. Arauzo-Carod (2011): “Locations and relocations: determinants, modelling, and interrelations,” The Annals of Regional Science, 47, 131–146.

Manniche, J. (2012): “Combinatorial knowledge dynamics: On the usefulness of the differentiated knowledge bases model,” European Planning Studies, 20, 1823–1841.

Martin, R. and J. Moodysson (2013): “Comparing knowledge bases: on the geography and organization of knowledge sourcing in the regional innovation system of Scania, Sweden,” European Urban and Regional Studies, 20, 170–187.

Max-Planck-Institute (2019): “Max Planck Institutes and Experts.” "https://www.mpg.de/institutes", [accessed on May 25 2019].

McFadden, D. (1973): “Conditional logit analysis of qualitative choice behavior,” .

Meijers, E. J., M. J. Burger, and M. M. Hoogerbrugge (2016): “Borrowing size in networks of cities: City size, network connectivity and metropolitan functions in Europe,” Papers in Regional Science, 95, 181–198.

Michelacci, C. and O. Silva (2007): “Why so many local entrepreneurs?” The Review of Economics and Statistics, 89, 615–633.

Moodysson, J. (2008): “Principles and practices of knowledge creation: On the organization of “buzz” and “pipelines” in life science communities,” Economic Geography, 84, 449–469.

Moretti, E. (2012): The new geography of jobs, Houghton Mifflin Harcourt.

Morley, C., J. Rosselló, and M. Santana-Gallego (2014): “Gravity models for tourism demand: theory and use,” Annals of Tourism Research, 48, 1–10.

Nguyen, C. Y., K. Sano, T. V. Tran, and T. T. Doan (2013): "Firm relocation patterns incorporating spatial interactions," The Annals of Regional Science, 50, 685–703.

North Data (2019): "Data Services," "https://www.northdata.com/_data", [accessed on January 20 2019].

Oakey, R., M. Kipling, and S. Wildgust (2001): "Clustering among firms in the non-broadcast visual communications (NBVC) sector," Regional Studies, 35, 401–414.

OEFW (2016): "Organisationen Einrichtungen in Forschung und Wissenschaft. Bundesbericht Forschung und Innovation 2016." .

Olivero, M. P. and Y. V. Yotov (2012): "Dynamic gravity: endogenous country size and asset accumulation," Canadian Journal of Economics/Revue canadienne d'économique, 45, 64–92.

OpenStreetMap (2018): "Germany," "https://download.geofabrik.de/europe/germany.html, [accessed on December 20 2018].

Ouwehand, W. M., F. G. van Oort, and N. Cortinovis (2022): "Spatial structure and productivity in European regions," Regional Studies, 56, 48–62.

Ozmen-Ertekin, D., K. Ozbay, and J. Holguin-Veras (2007): "Role of transportation accessibility in attracting new businesses to New Jersey," Journal of Urban Planning and Development, 133, 138–149.

Pan, Y., T. Conroy, A. Tsvetkova, and M. Kures (2020): "Incentives and firm migration: an interstate comparison approach," Economic Development Quarterly, 34, 140–153.

Parr, J. B. (2014): "The regional economy, spatial structure and regional urban systems," Regional Studies, 48, 1926–1938.

Phelps, N. A., R. J. Fallon, and C. Williams (2001): "Small firms, borrowed size and the urban-rural shift," Regional Studies, 35, 613–624.

Pijnenburg, K. and K. A. Kholodilin (2014): "Do regions with entrepreneurial neighbours perform better? A spatial econometric approach for German regions," Regional Studies, 48, 866–882.

Pina, K. and B. S. Tether (2016): "Towards understanding variety in knowledge intensive business services by distinguishing their knowledge bases," Research Policy, 45, 401–413.

Plum, O. and R. Hassink (2011): “Comparing knowledge networking in different knowledge bases in Germany,” Papers in Regional Science, 90, 355–371.

Porter, M. E. (1990): “The competitive Advantage of nations Macmillan press LTD,” .

Portes, R. and H. Rey (2005): “The determinants of cross-border equity flows,” Journal of International Economics, 65, 269–296.

Puga, D. (2010): “The magnitude and causes of agglomeration economies,” Journal of Regional Science, 50, 203–219.

Rammer, C. (2002): “Patente und Marken als Schutzmechanismen für Innovationen,” Studien zum deutschen Innovationssystem.

Rammer, C., J. Kinne, and K. Blind (2020): “Knowledge proximity and firm innovation: A microgeographic analysis for Berlin,” Urban Studies, 57, 996–1014.

Redding, S. J., D. M. Sturm, and N. Wolf (2011): “History and industry location: evidence from German airports,” Review of Economics and Statistics, 93, 814–831.

Rice, P., A. J. Venables, and E. Patacchini (2006): “Spatial determinants of productivity: Analysis for the regions of Great Britain,” Regional Science and Urban Economics, 36, 727–752.

Roche, M. P. (2020): “Taking innovation to the streets: microgeography, physical structure, and innovation,” Review of Economics and Statistics, 102, 912–928.

Romer, P. M. (1990): “Endogenous technological change,” Journal of Political Economy, 98, S71–S102.

Rosenthal, S. S. and W. C. Strange (2003): “Geography, industrial organization, and agglomeration,” Review of Economics and Statistics, 85, 377–393.

——— (2008): “The attenuation of human capital spillovers,” Journal of Urban Economics, 64, 373–389.

Rossi, F. and M. Dej (2020): “Where do firms relocate? Location optimisation within and between Polish metropolitan areas,” The Annals of Regional Science, 64, 615–640.

RWI and ImmobilienScout24 (2021): “RWI Real Estate Data- Apartments for Rent- suf. RWI-GEO-RED,” .

Saarenketo, S., K. Puumalainen, K. Kyläheiko, and O. Kuivalainen (2008): "Linking knowledge and internationalization in small and medium-sized enterprises in the ICT sector," Technovation, 28, 591–601.

Santos Silva, J. and S. Tenreyro (2022): "The log of gravity at 15," Portuguese Economic Journal, 1–15.

Schartinger, D., A. Schibany, and H. Gassler (2001): "Interactive relations between universities and firms: empirical evidence for Austria," The Journal of Technology Transfer, 26, 255–268.

Schumpeter, J. A. (1934): The theory of economic development – An inquiry into profits, capital, credit, interest, and the business cycle., Harvard University Press (repr. New Bunswick: Transaction, 1983).

Secundo, G., P. Del Vecchio, L. Simeone, and G. Schiuma (2020): "Creativity and stakeholders' engagement in open innovation: Design for knowledge translation in technology-intensive enterprises," Journal of Business Research, 119, 272–282.

Silva, J. S. and S. Tenreyro (2006): "The log of gravity," The Review of Economics and Statistics, 88, 641–658.

Sleutjes, B. and P. Beckers (2013): "Exploring the role of the neighbourhood in firm relocation: differences between stayers and movers," Journal of Housing and the Built Environment, 28, 417–433.

Smit, M. J. and H. L. De Groot (2013): "Stocking up: The role of temporal and spatial R&D stocks," European Planning Studies, 21, 637–665.

Solow, R. M. (1956): "A contribution to the theory of economic growth," The quarterly journal of economics, 70, 65–94.

Stam, E. (2007): "Why butterflies don 't leave: Locational behavior of entrepreneurial firms," Economic Geography, 83, 27–50.

Stuetzer, M., A. Brodeur, M. Obschonka, D. B. Audretsch, J. Rentfrow, J. Potter, and S. Gosling (2021): "A Golden Opportunity: The Gold Rush, Entrepreneurship and Culture," .

Taalbi, J. (2017): "What drives innovation? Evidence from economic history," Research Policy, 46, 1437–1453.

Tödtling, F. and M. Grillitsch (2015): “Does combinatorial knowledge lead to a better innovation performance of firms?” European Planning Studies, 23, 1741–1758.

Toner, P. (2010): “Innovation and vocational education,” The Economic and Labour Relations Review, 21, 75–98.

Trippl, M., F. Tödtling, and L. Lengauer (2009): “Knowledge sourcing beyond buzz and pipelines: evidence from the Vienna software sector,” Economic Geography, 85, 443–462.

Tura, T., V. Harmaakorpi, and S. Pekkola (2008): “Breaking inside the black box: towards a dynamic evaluation framework for regional innovative capability,” Science and Public Policy, 35, 733–744.

van Dijk, J. and P. Pellenbarg (2017): “Firm Migration,” International encyclopedia of geography: People, the Earth, Environment and Technology, 1.

van Dijk, J. and P. Pellenbarg (2017): Firm Migration, American Association of Geographers, 1–12, includes bibliographical references and index.

Van Dijk, J. and P. H. Pellenbarg (2000): “Firm relocation decisions in The Netherlands: An ordered logit approach,” Papers in Regional Science, 79, 191–219.

Van Oort, F., S. de Geus, and T. Dogaru (2015): “Related variety and regional economic growth in a cross-section of European urban regions,” European Planning Studies, 23, 1110–1127.

Van Oort, F. G. and O. A. Atzema (2004): “On the conceptualization of agglomeration economies: The case of new firm formation in the Dutch ICT sector,” The Annals of Regional Science, 38, 263–290.

Van Soest, D. P., S. Gerking, and F. G. Van Oort (2006): “Spatial impacts of agglomeration externalities,” Journal of Regional Science, 46, 881–899.

Van Winden, W., E. Braun, A. Otgaar, and J.-J. Witte (2014): Urban innovation systems: What makes them tick?, Routledge.

Van Wissen, L. (2000): “A micro-simulation model of firms: Applications of concepts of the demography of the firm,” Papers in Regional Science, 79, 111–134.

Venables, A. J. (2001): “Geography and international inequalities: the impact of new technologies,” Journal of Industry, Competition and Trade, 1, 135–159.

Volgmann, K. and A. Münter (2022): "Understanding metropolitan growth in German polycentric urban regions," Regional Studies, 56, 99–112.

Volgmann, K. and K. Rusche (2020): "The geography of borrowing size: Exploring spatial distributions for German urban regions," Tijdschrift voor economische en sociale geografie, 111, 60–79.

Weber, T., I. Bertschek, J. Ohnemus, and M. Ebert (2018): "Monitoring-Report Wirtschaft DIGITAL 2018," Tech. rep., ZEW-Gutachten und Forschungsberichte.

Weidenbach (2022): "Ranking der Top-50 Unternehmen weltweit nach ihrem Markenwert im Jahr 2021," "https://de.statista.com/statistik/daten/studie/162524/umfrage/markenwert-der-wertvollsten-unternehmen-weltweit/, [accessed on July 9 2022].

Weterings, A. and J. Knoben (2013): "Footloose: An analysis of the drivers of firm relocations over different distances," Papers in Regional Science, 92, 791–809.

Wooldridge, J. M. (2015): Introductory econometrics: A modern approach, Cengage learning.

Yi, Y. (2018): "Firm relocation and age-dependent reliance on agglomeration externalities," The Annals of Regional Science, 61, 439–456.

Yotov, Y. V., R. Piermartini, M. Larch, et al. (2016): An Advanced Guide to Trade Policy Analysis: The Structural Gravity Model, WTO iLibrary.

Appendix

A1 Appendix 1

Table A1.1: Variable Description and Sources of Regional Characteristics

Variable	**Description**	**Source**
Population density	Population per km^2	BBSR
Industry ratio	Employees at the place of work in industry (WZ 2008) per 100 inhabitants of working age	BBSR
Service ratio	Employees at work in the service sector (WZ 2008) per 100 inhabitants of working age	BBSR
Price index	Difference of the counties mean rent price (in percent) to the German mean price	RWI 2020
Price index change	Change in difference of the counties mean rent price (in percent) to the German mean rent price	RWI 2020
Gross income	Gross monthly earnings of employees in euros	BBSR
Firms per 1,000 inhabitants	INKAR; North Data	North Data (2019), BBSR
Universities	Number of universities and universities of applied science in counties	Hochschulkompass (2020)
Research Institutes	Number of publicly funded research institutes in district and institutes of the four biggest research organizations in Germany	Fraunhofer, Helmholtz, Leibniz, Max-Planck-Institutes, Forschungseinrichtungen des Bundes und der Länder
Metropole	Cities with more than 500 000 inhabitants in 2017: Hamburg, Berlin, Munich, Stuttgart, Cologne, Frankfurt am Main, Düsseldorf, Dortmund, Essen	BBSR
Distance	Linear distance in kilometer (as the crow flies)	Own calculation with R 'geocode' command

Notes: The table gives detailed information on the variables and its sources.

Table A1.2: Correlation Table on Regional Characteristics

	Bilateral flow	Industrial ratio	Service ratio	Price index	Price index change	Firms per 1,000 inhabitants	Research institutes	Universities	Firm birth (ln)	Gross income (ln)	Population density (ln)
Bilateral flow	1	-0.02	0.05	0.08	0.04	0.10	0.10	0.11	0.09	0.06	0.06
Industrial ratio	-0.02	1	0.12	-0.06	0.14	-0.06	-0.17	-0.11	-0.19	0.39	0.01
Service ratio	0.05	0.12	1	0.44	0.30	0.57	0.42	0.44	0.32	0.45	0.66
Price index	0.08	-0.06	0.44	1	0.31	0.70	0.41	0.41	0.65	0.58	0.52
Price index change	0.04	0.14	0.30	0.31	1	0.49	0.20	0.19	0.13	0.48	0.17
Firms per 1,000 inhabitants	0.10	-0.06	0.57	0.70	0.49	1	0.50	0.48	0.61	0.63	0.51
Research institutes	0.10	-0.17	0.42	0.41	0.20	0.50	1	0.80	0.57	0.25	0.45
Universities	0.11	-0.11	0.44	0.41	0.19	0.48	0.80	1	0.63	0.31	0.51
Firm birth (ln)	0.09	-0.19	0.32	0.65	0.13	0.61	0.57	0.63	1	0.44	0.56
Gross income (ln)	0.06	0.39	0.45	0.58	0.48	0.63	0.25	0.31	0.44	1	0.54
Population density (ln)	0.06	0.01	0.66	0.52	0.17	0.51	0.45	0.51	0.56	0.54	1

Notes: This table displays the correlation between all control variables.

Table A1.3: Variance Inflation Factor for Regression 1.4 (Pooled OLS total firm birth)

Population density	2.769
Industrial ratio	1.843
Service ratio	2.626
Gross income	3.798
Firms per 1,000 inhabitants	4.223
Price index	2.232
Price Index change	1.731
Metropole	2.080
Neighbour is metro	1.209
Lag(firmbirth)	3.263
Universities	3.918
Research Institutes	3.068

Notes: The variance inflation factor is a measure of the amount of multicollinearity in a set of multiple regression variables. A large variance inflation factor (>10) on an independent variable indicates a highly collinear relationship to the other variables.

Table A1.4: Gravity Model (2) Robustness

	PPML without Munich -Pair (1.1)		Negative Binomial (1.2)		PPML only positive (1.3)		Negative binomial only positive (1.4)		OLS only positive (1.5)	
population density (ln) Origin	1.63213	(1.3885)	2.83420*	(1.42778)	0.02343	(0.83008)	0.07375	(0.97812)	0.08120	(0.41348)
Population density (ln) Dest.	0.71251	(0.98430)	-0.48506	(1.37770)	0.05778	(0.79220)	0.10875	(0.93831)	0.27016	(0.39645)
Industrial ratio Origin	0.01880	(0.01519)	0.01525	(0.02163)	-0.01548	(0.01075)	-0.01530	(0.01451)	-0.00856	(0.00546)
Industrial ratio Destination	0.00919	(0.01583)	0.00304	(0.02129)	0.00439	(0.01091)	0.00406	(0.01462)	0.00501	(0.00582)
Service ratio Origin	0.02295	(0.01209)	0.01360	(0.01674)	0.00729	(0.00868)	0.00710	(0.01125)	0.00473	(0.00457)
Service ratio Destination	0.00711	(0.01204)	-0.00410	(0.01664)	-0.00175	(0.00959)	-0.00163	(0.01134)	-0.00010	(0.00471)
Gross income (ln) Origin	-1.45613	(0.82506)	-2.34002*	(1.14170)	-0.45356	(0.57452)	-0.44300	(0.77407)	-0.25914	(0.32720)
Gross income (ln) Destination	0.70904	(0.81127)	0.58105	(1.12996)	0.34709	(0.59761)	0.33541	(0.76130)	-0.01262	(0.31153)
Price index Origin	-0.00139	(0.01023)	0.01339	(0.01518)	-0.02030*	(0.01005)	-0.01980*	(0.00876)	-0.00843	(0.00477)
Price index Destination	0.02161*	(0.01031)	0.01787	(0.01486)	-0.00237	(0.01068)	-0.00297	(0.00881)	-0.00381	(0.00466)
Price index change Origin	-0.00165	(0.00899)	-0.01426	(0.01336)	0.01762	(0.00970)	0.01699*	(0.00743)	0.00718	(0.00441)
Price index change Destination	-0.02168*	(0.00925)	-0.01845	(0.01311)	0.00213	(0.01009)	0.00257	(0.00748)	0.00298	(0.00431)
Firm birth (ln) Origin	-0.08206	(0.04667)	-0.02441	(0.06051)	-0.00115	(0.02821)	-0.00135	(0.04651)	-0.00409	(0.01575)
Firm birth (ln) Destination	0.02143	(0.04442)	-0.00645	(0.05785)	0.00596	(0.02775)	0.00595	(0.04428)	-0.00413	(0.01485)
Firms per 1,000inhabitants Origin	-0.08646	(0.08722)	-0.08917	(0.12763)	0.01308	(0.08749)	0.00675	(0.07311)	-0.01731	(0.04150)
Firms per 1,000inhabitants Dest.	0.25127**	(0.08703)	0.33920**	(0.12973)	0.10950	(0.10334)	0.09478	(0.07114)	0.02160	(0.04525)
Research Institutes Origin	-0.01354	(0.04253)	-0.08155	(0.07059)	-0.02616	(0.03787)	-0.02441	(0.03426)	-0.01744	(0.02027)
Research Institutes Destination	-0.06343	(0.04540)	-0.13990	(0.07578)	0.00495	(0.04893)	0.00543	(0.03692)	0.00529	(0.02327)
Universities Origin	0.06766*	(0.03000)	0.05790	(0.05153)	0.06622*	(0.02720)	0.06647**	(0.02546)	0.03583*	(0.01508)
Universities Destination	0.00625	(0.03309)	0.02285	(0.05370)	0.01333	(0.03513)	0.01541	(0.02600)	0.00503	(0.01690)
Neighbour county	1.19925***	(0.04095)	1.40058***	(0.05799)	0.28267***	(0.02196)	0.28259***	(0.03116)	0.19787***	(0.01400)
Distance (ln)	-1.20744***	(0.01584)	-1.40641***	(0.02122)	-0.26496***	(0.01533)	-0.25571***	(0.01253)	-0.11935***	(0.00617)
Num. obs.	1,432,800		1,432,818		10,108		10,108		10,108	

Notes: Dependent variable is M_{ijt}, significant levels: ***p < 0.001; **p < 0.01; *p < 0.05; time, origin and destination fixed effects included, standard errors in parentheses; dependent variable as log-link.

Table A1.5: Gravity Model (3) (Difference Approach) Robustness

	PPML without Munich-Pair (2.1)	Negative binomial (2.2)	PPML only positive (2.3)	Negative binomial only positive (2.4)	OLS only positive (2.5)
Population density	0.00011	0.00023	-0.00032	-0.00029	-0.00004
	(0.00027)	(0.00045)	(0.00032)	(0.00022)	(0.00001)
Industrial ratio	-0.00108	0.00300	0.00530	0.00545	0.00591
	(0.01077)	(0.01524)	(0.00677)	(0.01028)	(0.00005)
Service ratio	-0.00582	0.00104	-0.00754	-0.00697	-0.00292
	(0.00778)	(0.01114)	(0.00636)	(0.00726)	(0.00004)
Gross income	0.00026	0.00032	0.00014	0.00013	0.00004
	(0.00022)	(0.00030)	(0.00015)	(0.00020)	(0.00000)
Price index	0.01850*	0.01122	0.00762	0.00734	0.00201*
	(0.00866)	(0.01143)	(0.00879)	(0.00718)	(0.00004)
Price index change	-0.01624*	-0.01086	-0.00629	-0.00599	-0.00176*
	(0.00778)	(0.01015)	(0.00835)	(0.00624)	(0.00004)
Firm birth	-0.00029	-0.00027	0.00002	0.00002	0.00004
	(0.00032)	(0.00055)	(0.00034)	(0.00025)	(0.00001)
Firms per 1,000 inhabitants	0.17277**	0.15518	0.05985	0.05381	0.01784
	(0.06527)	(0.09082)	(0.08113)	(0.05042)	(0.00069)
Research Institutes	-0.02444	-0.02914	0.02234	0.02103	0.01261
	(0.03320)	(0.05313)	(0.03264)	(0.02482)	(0.00071)
Universities	-0.01512	-0.00456	-0.01982	-0.01954	-0.01658
	(0.02654)	(0.04770)	(0.02329)	(0.02254)	(0.00065)
Neighbour county	1.19930***	1.40335***	0.28140***	0.28118***	0.19704***
	(0.04097)	(0.05831)	(0.02209)	(0.03117)	(0.00263)
Distance (ln)	-1.20746***	-1.40692***	-0.26403***	-0.25460***	-0.11926***
	(0.01584)	(0.02131)	(0.01550)	(0.01253)	(0.00020)
Num. obs.	1,432,800	1,432,818	10,108	10,108	10,108

Notes: significant levels: ***p $<$ 0.001; **p $<$ 0.01; *p $<$ 0.05; time, origin and destination fixed effects included, standard errors in parentheses; dependent variable as log-link.

Table A1.6: Regression Multilateral Resistance

	(1)
Neighbour County	1.21401***
	(0.04061)
Distance (ln)	-1.16234***
	(0.01617)
Time-County fixed effect	Yes
Num. obs.	770,120

Notes: The Table shows Model (2) (PPML Gravity Model) with county-time fixed effects. Dependent variable is M_{ijt}. Significant levels: ***p $<$ 0.001; **p $<$ 0.01; *p $<$ 0.05.

Table A1.7: Binary Regressions: Push and Pull Factors

	push factors			**pull factors**		
	OLS	logistic	probit	OLS	logistic	probit
	(1.1)	(1.2)	(1.3)	(2.1)	(2.2)	(2.3)
Firmage	-0.003***	-0.117***	-0.050***	0.001***	0.025***	0.011***
	(0.0001)	(0.004)	(0.002)	(0.0001)	(0.003)	(0.001)
Population density	-0.00000	-0.0001	-0.00003	-0.00001*	-0.0003*	-0.0001*
	(0.00000)	(0.0002)	(0.0001)	(0.00000)	(0.0002)	(0.0001)
Industrial ratio	0.0002	0.002	0.001	0.0003	0.009	0.004
	(0.0002)	(0.006)	(0.003)	(0.0002)	(0.008)	(0.003)
Service ratio	-0.0003	-0.011	-0.005	0.0001	0.004	0.001
	(0.0003)	(0.008)	(0.004)	(0.0003)	(0.010)	(0.004)
Gross income	-0.003	-0.048	-0.030	0.004	0.111	0.052
	(0.008)	(0.232)	(0.102)	(0.007)	(0.262)	(0.112)
Stock of firms	0.00000**	0.00002	0.00001	-0.00000	-0.00002	-0.00001
	(0.00000)	(0.00003)	(0.00001)	(0.00000)	(0.00003)	(0.00001)
Firms per 1,000 inhabitants	-0.003	-0.003	-0.006	0.004*	0.112*	0.047*
	(0.002)	(0.055)	(0.025)	(0.002)	(0.066)	(0.029)
Price Index	0.0002	0.001	0.001	0.0003	0.011	0.005
	(0.0002)	(0.007)	(0.003)	(0.0002)	(0.008)	(0.004)
Price Index change	-0.0002	-0.001	-0.001	-0.0003*	-0.011	-0.005
	(0.0002)	(0.006)	(0.003)	(0.0002)	(0.007)	(0.003)
Research Institutes	-0.001	-0.034	-0.015	-0.001	-0.025	-0.012
	(0.001)	(0.032)	(0.014)	(0.001)	(0.037)	(0.016)
Universities	-0.0005	-0.00003	-0.0002	0.0001	0.007	0.002
	(0.001)	(0.029)	(0.013)	(0.001)	(0.034)	(0.014)
Firm birth (lag)	-0.00001	-0.0002	-0.0001	-0.00001	-0.0002	-0.0001
	(0.00001)	(0.0003)	(0.0001)	(0.00001)	(0.0004)	(0.0002)
metropole	0.006	0.142	0.060	0.006*	0.289**	0.121**
	(0.004)	(0.137)	(0.058)	(0.003)	(0.146)	(0.061)
Neighbour is metro	-0.005	-0.085	-0.039	-0.018	-0.554*	-0.244*
	(0.012)	(0.317)	(0.141)	(0.011)	(0.332)	(0.146)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	552,384	552,384	552,384	552,384	552,384	552,384
Adjusted R^2	0.006			0.005		
Log Likelihood		-75,148.870	-75,157.100		-67,201.860	-67,199.900

Notes: Significant levels: ***p < 0.001; **p < 0.01; *p < 0.05, cluster robust standard errors in parentheses.

Table A1.8: Gravity Model (2) – subsample ICT firms only

	PPML	NB	OLS only positive	PPML only positive	NB only positive
	(1)	(2)	(3)	(4)	(5)
population density (ln) Origin	0.45225	3.11375	0.18712	-0.85236	-0.81696
	(1.93014)	(2.28018)	(0.81547)	(1.38927)	(1.38434)
population density (ln) Destination	-2.96359	0.05174	-0.97619	-2.15182	-2.13519
	(1.98166)	(2.35171)	(0.69926)	(1.15194)	(1.14870)
industrial ratio Origin	-0.00094	-0.00517	-0.01042	-0.01219	-0.01224
	(0.02507)	(0.02936)	(0.00807)	(0.01450)	(0.01443)
industrial ratio Destination	-0.01794	-0.02089	-0.00048	0.00648	0.00623
	(0.02467)	(0.02818)	(0.00940)	(0.01507)	(0.01502)
service ratio Origin	0.01564	0.02259	0.00433	0.00166	0.00169
	(0.01951)	(0.02187)	(0.00783)	(0.01287)	(0.01282)
service ratio Destination	-0.03691	-0.01703	-0.00423	-0.00326	-0.00338
	(0.02126)	(0.02405)	(0.00756)	(0.01265)	(0.01262)
gross income (ln) Origin	2.20674	-1.18291	0.85534	1.45821	1.45335
	(1.56482)	(1.78579)	(0.55553)	(0.90124)	(0.89835)
gross income (ln) Destination	-2.35208	-2.79974	0.49265	0.30016	0.30794
	(1.51402)	(1.75435)	(0.53697)	(0.86886)	(0.86476)
Price index Origin	-0.05542**	-0.02861	-0.01773*	-0.02699	-0.02693
	(0.01994)	(0.02478)	(0.00849)	(0.01424)	(0.01418)
Price index Destination	0.00743	0.03359	-0.00775	-0.00985	-0.01000
	(0.02021)	(0.02450)	(0.00874)	(0.01422)	(0.01420)
Price index change origin	0.04724**	0.02281	0.01469	0.02419	0.02407
	(0.01735)	(0.02172)	(0.00779)	(0.01325)	(0.01319)
Price index change Destination	-0.01082	-0.03464	0.00625	0.00778	0.00791
	(0.01766)	(0.02132)	(0.00793)	(0.01286)	(0.01285)
Firmbirth ICT (ln) Origin	-0.17063**	-0.19524**	-0.03447	-0.08933**	-0.08865**
	(0.05925)	(0.06770)	(0.01854)	(0.02839)	(0.02830)
Firmbirth ICT (ln) Destination	-0.05265	-0.03699	-0.00399	0.00554	0.00544
	(0.05720)	(0.06506)	(0.01927)	(0.03061)	(0.03046)
Firms per 1,000 inhabitants Origin	-0.35879	-0.33030	-0.02109	0.18684	0.18257
	(0.26393)	(0.32578)	(0.12798)	(0.20201)	(0.20182)
Firms per 1,000 inhabitants Destination	0.67789*	0.90800**	0.05192	0.27037	0.26430
	(0.26916)	(0.33666)	(0.12507)	(0.20564)	(0.20554)
Research Institutes Origin	-0.06812	-0.21143*	-0.00152	-0.00399	-0.00408
	(0.06999)	(0.09443)	(0.02982)	(0.04769)	(0.04744)
Research Institutes Destination	-0.11007	-0.25765**	0.02129	0.01338	0.01305
	(0.07554)	(0.09871)	(0.03279)	(0.05508)	(0.05478)
Universities Origin	-0.05135	0.06329	0.00603	-0.01576	-0.01528
	(0.06088)	(0.08726)	(0.02586)	(0.03450)	(0.03443)
Universities Destination	-0.06391	-0.02075	0.01330	0.01424	0.01462
	(0.06644)	(0.09226)	(0.02584)	(0.03859)	(0.03853)
Neighbour County	1.06446***	1.19518***	0.11101***	0.13765***	0.13805***
	(0.06724)	(0.07392)	(0.01930)	(0.02835)	(0.02825)
Distance (ln)	-1.25790***	-1.44787***	-0.09906***	-0.21064***	-0.20901***
	(0.02828)	(0.03163)	(0.01008)	(0.01746)	(0.01749)
Num. obs.	820,890	820,890	3,537	3,537	3,537

Notes: significance levels: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. Year, origin and destination fixed effects included; standard errors in parentheses; variables coded as differences between origin and destination; firm birth and firms per 1,000 inhabitants refer to digital firms; dependent variable as log-link. The subsample contains only ICT firms selected via NACE code.

Table A1.9: Gravity Model (3) (Difference Approach) – subsample ICT Firms only

	PPML	NB	OLS only positive	PPML only positive	NB only positive
	(1)	(2)	(3)	(4)	(5)
Population density	-0.00057	-0.00078	-0.00046	-0.00089*	-0.00088*
	(0.00055)	(0.00072)	(0.00026)	(0.00044)	(0.00044)
Industrial ratio	-0.00539	-0.00536	0.00218	0.00387	0.00384
	(0.01769)	(0.02093)	(0.00575)	(0.00940)	(0.00936)
Service ratio	-0.02315	-0.01935	-0.00578	-0.00688	-0.00682
	(0.01379)	(0.01578)	(0.00508)	(0.00922)	(0.00917)
Gross income	-0.00093*	-0.00060	0.00004	-0.00005	-0.00005
	(0.00037)	(0.00042)	(0.00013)	(0.00021)	(0.00021)
Price index	0.02802	0.02856	0.00175	0.00300	0.00292
	(0.01548)	(0.01783)	(0.00585)	(0.00995)	(0.00989)
Price Index change	-0.02508	-0.02570	-0.00108	-0.00285	-0.00274
	(0.01374)	(0.01579)	(0.00544)	(0.00951)	(0.00946)
Firm birth	0.00084	0.00036	0.00040	0.00142	0.00141
	(0.00100)	(0.00139)	(0.00056)	(0.00099)	(0.00098)
Firms per 1,000 inhabitants	0.58614**	0.71947**	0.03783	0.05266	0.04983
	(0.20703)	(0.25372)	(0.10024)	(0.18772)	(0.18659)
Research Institutes	-0.02493	-0.01936	0.01151	0.01459	0.01421
	(0.05147)	(0.07030)	(0.02285)	(0.04223)	(0.04187)
Universities	0.00533	-0.02566	0.01019	0.02640	0.02616
	(0.04681)	(0.06261)	(0.01850)	(0.02498)	(0.02493)
Neighbour county	1.06457***	1.19435***	0.11181***	0.14177***	0.14203***
	(0.06736)	(0.07390)	(0.01929)	(0.02831)	(0.02821)
Distance (ln)	-1.25799***	-1.44694***	-0.09842***	-0.20645***	-0.20497***
	(0.02835)	(0.03163)	(0.01001)	(0.01790)	(0.01789)
Num. obs.	820,890	820,890	3,537	3,537	3,537

Notes: Year, origin and destination fixed effects included; standard errors in parentheses; variables coded as differences between origin and destination; firm birth and firms per 1,000 inhabitants refer to digital firms; dependent variable as log-link. The subsample contains only ICT firms selected via NACE code. Significant levels: ***p < 0.001; **p < 0.01; *p < 0.05.

Table A1.10: Gravity Model (2) – NUTS 2 Aggregate

	OLS (1)	PPML (2)	NB (3)	OLS OP (4)	PPML OP (5)	NB OP (6)
population density (ln) Origin	-0.00499	-0.76959	-0.32974	-2.06923	-2.54100	-2.38491
	(0.46466)	(2.68611)	(2.63817)	(1.59631)	(2.29354)	(2.18228)
population density (ln) Destination	0.16195	1.96055	3.55779	1.18215	0.50148	0.75347
	(0.46319)	(2.58177)	(2.49504)	(1.61529)	(2.22609)	(2.12256)
industrial ratio Origin	0.00075	0.01130	0.00717	-0.01894	-0.00335	-0.00694
	(0.00781)	(0.03929)	(0.04025)	(0.02425)	(0.03122)	(0.03014)
industrial ratio Destination	-0.00468	0.00407	0.00121	0.01950	0.02930	0.02877
	(0.00793)	(0.03797)	(0.03780)	(0.02242)	(0.03020)	(0.02901)
service ratio Origin	-0.00351	0.00548	0.04261	-0.00670	-0.05766	-0.04401
	(0.01047)	(0.05473)	(0.05441)	(0.03363)	(0.04622)	(0.04440)
service ratio Destination	0.00143	0.03576	0.04383	0.01014	0.00950	0.01179
	(0.01053)	(0.05310)	(0.05309)	(0.03312)	(0.04554)	(0.04358)
gross income (ln) Origin	-1.43326***	-3.93634	-5.42891**	-3.42787**	-3.80443*	-4.19796*
	(0.41605)	(2.02906)	(2.09358)	(1.26540)	(1.68390)	(1.63587)
gross income (ln) Destination	-1.05697*	0.00903	-0.52292	0.70540	-0.66257	-0.87978
	(0.41537)	(2.02828)	(2.04743)	(1.32529)	(1.75547)	(1.69538)
Price index Origin	-0.01080	0.00346	-0.00776	-0.02134	-0.01633	-0.02130
	(0.00654)	(0.02036)	(0.02142)	(0.01475)	(0.01858)	(0.01801)
Price index Destination	-0.00913	-0.00655	-0.00648	-0.01750	-0.02452	-0.02427
	(0.00657)	(0.02093)	(0.02130)	(0.01530)	(0.01951)	(0.01886)
Price index change origin	0.00969	0.00015	0.01904	0.02273	0.01036	0.01747
	(0.00631)	(0.02148)	(0.02233)	(0.01553)	(0.01965)	(0.01901)
Price index change Destination	0.00773	-0.00884	-0.00405	-0.00064	-0.00214	-0.00058
	(0.00622)	(0.02257)	(0.02233)	(0.01594)	(0.02081)	(0.02012)
Firmbirth (ln) Origin	-0.02450	0.09521	-0.10791	0.12112	0.31933	0.26120
	(0.03729)	(0.20348)	(0.20913)	(0.11539)	(0.16293)	(0.15754)
Firmbirth (ln) Destination	0.05162	0.40652*	0.38658	0.29857*	0.40437*	0.35582*
	(0.03755)	(0.19797)	(0.19940)	(0.12178)	(0.16959)	(0.16160)
Firms per 1,000 inhabitants Origin	0.19808**	0.09573	-0.13053	0.11375	0.40739*	0.31195
	(0.06525)	(0.22292)	(0.23002)	(0.14808)	(0.19203)	(0.18529)
Firms per 1,000 inhabitants Destination	0.10075	0.01102	-0.19011	0.26824	0.37976	0.34314
	(0.06559)	(0.22888)	(0.23190)	(0.15585)	(0.20351)	(0.19700)
Research Institutes Origin	0.02181	-0.04041	-0.02375	0.01132	-0.02360	-0.02103
	(0.01180)	(0.03880)	(0.03816)	(0.02653)	(0.03424)	(0.03321)
Research Institutes Destination	0.01547	-0.02870	-0.03115	-0.01964	-0.01476	-0.01450
	(0.01144)	(0.03652)	(0.03788)	(0.02617)	(0.03056)	(0.03027)
Universities Origin	0.01778*	0.00372	-0.01385	0.00689	0.05244*	0.04349
	(0.00884)	(0.03018)	(0.03093)	(0.01915)	(0.02456)	(0.02384)
Universities Destination	0.04778***	0.07830*	0.08383**	0.02584	0.08376**	0.07074**
	(0.00877)	(0.03157)	(0.03172)	(0.02049)	(0.02576)	(0.02482)
Neighbour County	0.28837***	0.62898***	0.55653***	0.31860***	0.56386***	0.52520***
	(0.02829)	(0.06108)	(0.06145)	(0.04237)	(0.04543)	(0.04531)
Distance (ln)	-0.31075***	-0.79693***	-1.01996***	-0.44925***	-0.47760***	-0.48146***
	(0.01002)	(0.03365)	(0.03069)	(0.01794)	(0.01846)	(0.01856)
Num. obs.	12,654	11,340	11,340	3,291	3,291	3,291

Notes: Year, origin and destination fixed effects included; standard errors in parentheses; variables coded as differences between origin and destination; firm birth and firms per 1,000 inhabitants refer to digital firms; dependent variable as log-link. OP is only positive flows. The subsample contains only ICT firms selected via NACE code. Significant levels: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Table A1.11: Gravity Model (3) (Difference Approach) – NUTS 2 Aggregate

	OLS (1)	PPML (2)	NB (3)	OLS OP (4)	PPML OP (5)	NB OP (6)
Population density	0.00001	-0.00006	-0.00016	0.00028	0.00015	0.00019
	(0.00029)	(0.00053)	(0.00052)	(0.00042)	(0.00052)	(0.00050)
industrial ratio	-0.00369	-0.00858	-0.01107	0.02044	0.01789	0.01922
	(0.00539)	(0.02587)	(0.02643)	(0.01529)	(0.02643)	(0.01963)
Service ratio	0.00165	0.01063	-0.00810	0.01457	0.03731	0.03397
	(0.00662)	(0.03062)	(0.03002)	(0.01971)	(0.03002)	(0.02568)
Gross income	0.00007	0.00028	0.00061	0.00015	-0.00007	-0.00004
	(0.00011)	(0.00055)	(0.00054)	(0.00033)	(0.00054)	(0.00045)
Price index	0.00082	-0.00289	0.00136	-0.00790	-0.00958	-0.00856
	(0.00214)	(0.00930)	(0.00896)	(0.00583)	(0.00896)	(0.00753)
Price Index change	-0.00085	-0.00304	-0.00786	0.00084	0.00210	0.00121
	(0.00198)	(0.00943)	(0.00908)	(0.00584)	(0.00908)	(0.00772)
Firm birth	0.00002	-0.00024	-0.00008	-0.00006	-0.00021	-0.00019
	(0.00014)	(0.00029)	(0.00031)	(0.00024)	(0.00031)	(0.00028)
Firms per 1,000 inhabitants	-0.04450	0.01866	0.03418	0.00611	-0.01850	-0.00912
	(0.04288)	(0.13466)	(0.13197)	(0.08725)	(0.13197)	(0.10458)
Research Institutes	-0.00443	0.00703	-0.00484	-0.01172	0.00997	0.00823
	(0.00811)	(0.02796)	(0.02749)	(0.01855)	(0.02749)	(0.02253)
Universities	0.01505*	0.04858*	0.05551*	0.01740	0.02651*	0.02348
	(0.00614)	(0.02206)	(0.02157)	(0.01363)	(0.02157)	(0.01754)
Neighbour county	0.28837***	0.62907***	0.55606***	0.31860***	0.56450***	0.51797***
	(0.02862)	(0.06131)	(0.06138)	(0.04270)	(0.06138)	(0.04664)
Distance (ln)	-0.31075***	-0.79681***	-1.01876***	-0.44900***	-0.47800***	-0.48345***
	(0.01035)	(0.03369)	(0.03073)	(0.01844)	(0.03073)	(0.01965)
Num. obs.	12,654	11,340	11,340	3,291	3,291	3,291

Notes: Year, origin and destination fixed effects included; standard errors in parentheses; variables coded as differences between origin and destination; firm birth and firms per 1,000 inhabitants refer to digital firms; dependent variable as log-link. OP is only positive flows. Significant levels: ***p < 0.001; **p < 0.01; *p < 0.05.

A2 Appendix 2

Table A2.1: Summary Statistics of (control) variables

Statistic	N	Mean	St. Dev.	Min	Max
year	46,598	2,006	6.633	1,995	2,017
Share of firm birth	46,534	1.069	3.046	0.000	100.000
University Dummy	46,598	0.036	0.186	0	1
Population: 10,000 inhabitants	46,598	1.694	2.015	0.130	38.000
Share of firm birth next Center	46,534	44.470	26.267	0.000	100.000
Universities next Center	46,598	5.266	5.053	0	37
Population: 10,000 inhabitants next Center	46,598	50.020	64.195	9.810	361.349
Distance to next center	46,598	22.455	10.964	3.510	114.710

Notes: The Table shows the summary statistics of the control variables employed in the empirical analysis. Note that for the share of firm birth, there are two observations where 100 % of the firms were set up in one municipality, which is each driven by only one firm. This is Seesen in 2000 and Lutter am Barenberge in 2011 (both in Region Salzgitter).

Table A2.2: Summary Statistics of Distance Quintiles (in km) for MURs

Distance quintile	min	q1	median	mean	q3	max
1	4	10	13	13	15	42
2	18	21	23	24	26	55
3	29	31	34	35	38	62
4	41	43	47	47	50	61
5	11	11	63	63	115	115

Notes: The Table shows Summary Statistics on the Distance Quintiles for monocentric urban regions. It shows the relative measure, where the each region is divided into five Quintiles.

Table A2.3: Summary Statistics of Distance Quintiles (in km) for PURs

Distance quintile	min	q1	median	mean	q3	max
1	5	11	13	13	16	40
2	14	21	23	24	26	38
3	29	31	34	34	38	41
4	17	44	45	46	48	57
5	60	63	65	65	69	70

Notes: The Table shows Summary Statistics on the Distance Quintiles for polycentric urban regions. It shows the relative measure, where the each region is divided into five Quintiles.

Table A2.4: Correlation Table

	Population: 10,000 inhabitants	Population: 10,000 inhabitants next Center	University (dummy)	Universities next Center	Share of firm birth next Center	Share of firm birth
Population: 10,000 inhabitants	1.000	-0.045	0.410	-0.030	-0.191	0.220
Population: 10,000 inhabitants next Center	-0.045	1.000	0.009	0.885	0.441	-0.124
University (dummy)	0.410	0.009	1.000	0.011	-0.059	0.135
Universities next Center	-0.030	0.885	0.011	1.000	0.434	-0.113
Share of firm birth next Center	-0.191	0.441	-0.059	0.434	1.000	-0.003
Share of firm birth	0.220	-0.124	0.135	-0.113	-0.003	1.000

Notes: This table displays the correlation between all control variables.

Table A2.5: Robustness of the Interaction Terms

	Full (1)	MUR (2)	PUR (3)	MUR (4)	PUR (5)
Population: 10,000 inhabitants	0.380***	0.880***	0.210***	0.880***	0.210***
	(0.013)	(0.025)	(0.008)	(0.026)	(0.008)
University Dummy	0.950***	0.880***	0.300***	0.870***	0.290***
	(0.110)	(0.180)	(0.069)	(0.180)	(0.069)
Population: 10,000 inhabitants next Center	0.008***	0.055***	-0.0003	0.053***	-0.0002
	(0.001)	(0.006)	(0.0003)	(0.006)	(0.0003)
Universities next Center	0.040***	0.026**	-0.003	0.025**	-0.003
	(0.005)	(0.011)	(0.003)	(0.011)	(0.003)
Absolute firm birth Region	-0.0002***	-0.001***	-0.0001	-0.001***	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Share of firm birth next Center	-0.039***	-0.048***	0.004***	-0.048***	0.003***
	(0.003)	(0.004)	(0.001)	(0.004)	(0.001)
Distance to next center	-0.009***	-0.009***	-0.006***		
	(0.001)	(0.002)	(0.0005)		
dist_quintile2				-0.280***	-0.180***
				(0.042)	(0.018)
dist_quintile3				-0.150**	-0.190***
				(0.064)	(0.017)
dist_quintile4				-0.094*	-0.170***
				(0.057)	(0.017)
dist_quintile5				-0.800***	-0.260***
				(0.200)	(0.021)
Time FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Observations	46,534	31,722	14,812	31,722	14,812
Adjusted R^2	0.170	0.190	0.380	0.190	0.380

Notes: Dependent variable is the share of firm birth in the municipalities. Significant levels: ***p < 0.001; **p < 0.01; *p < 0.05.

Table A2.6: Robustness Absolute Distance Quintiles

	MUR (1)	PUR (2)
Population: 10,000 inhabitants	0.880***	0.210***
	(0.025)	(0.008)
University Dummy	0.860***	0.310***
	(0.180)	(0.070)
Population: 10,000 inhabitants next Center	0.054***	-0.0003
	(0.007)	(0.0003)
Universities next Center	0.025**	-0.003
	(0.011)	(0.003)
Absolute firm birth Region	-0.001***	-0.0001
	(0.0001)	(0.0001)
Share of firm birth next Center	-0.048***	0.004***
	(0.004)	(0.001)
dist_quintile_absolut2	-0.480***	-0.072***
	(0.063)	(0.026)
dist_quintile_absolut3	-0.460***	-0.078**
	(0.069)	(0.033)
dist_quintile_absolut4	-0.550***	-0.220***
	(0.067)	(0.024)
dist_quintile_absolut5	-0.430***	-0.180***
	(0.072)	(0.023)
Time FE	Yes	Yes
Region FE	Yes	Yes
Observations	31,722	14,812
Adjusted R^2	0.190	0.380

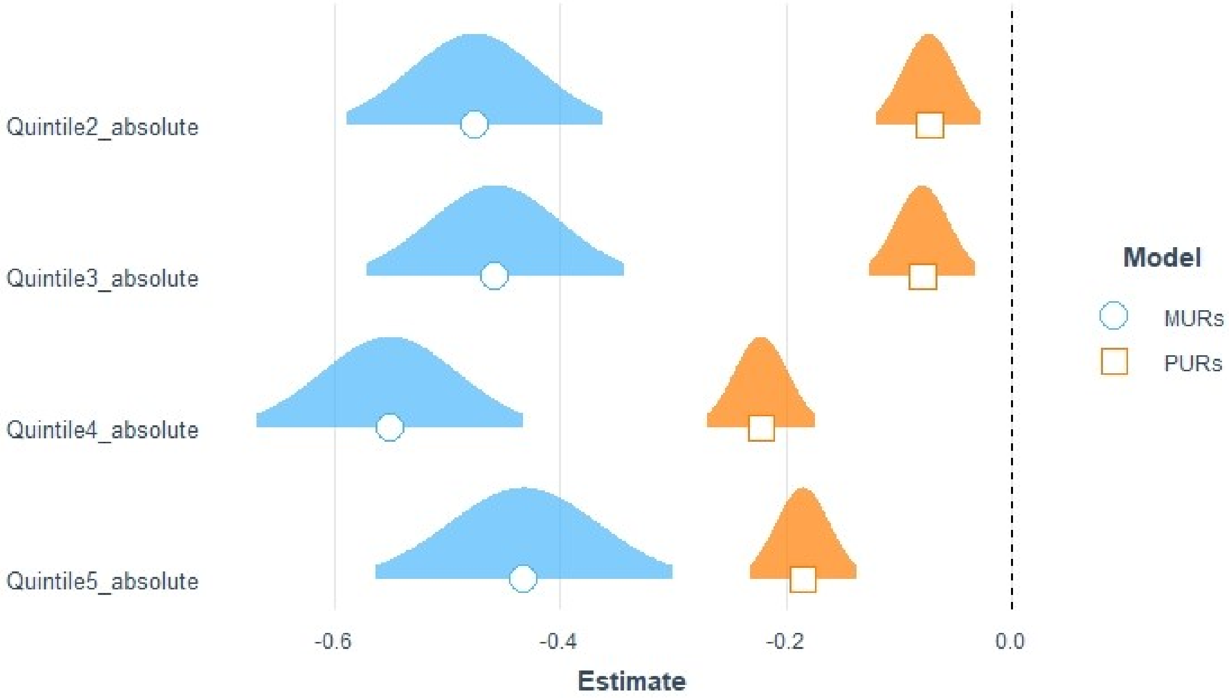
Notes: Dependent variable is the share of firm birth in the municipalities.
Significant levels: ***p < 0.001; **p < 0.01; *p < 0.05.

Table A2.7: Robustness Estimation including MUR Dummy

	(1)	(2)	(3)
Population: 10,000 inhabitants	0.380***	0.380***	0.380***
	(0.013)	(0.013)	(0.013)
University Dummy	0.950***	0.940***	0.940***
	(0.110)	(0.110)	(0.110)
Population: 10,000 inhabitants next Center	0.008***	0.008***	0.009***
	(0.001)	(0.001)	(0.001)
Universities next Center	0.040***	0.040***	0.041***
	(0.005)	(0.005)	(0.005)
Absolute firm birth Region	-0.0002***	-0.0002***	-0.0002***
	(0.0001)	(0.0001)	(0.0001)
Share of firm birth next Center	-0.039***	-0.039***	-0.039***
	(0.003)	(0.003)	(0.003)
Distance to next center	-0.009***		
	(0.001)		
dist_quintile2		-0.240***	-0.270***
		(0.031)	(0.061)
dist_quintile3		-0.190***	-0.180**
		(0.041)	(0.073)
dist_quintile4		-0.190***	-0.110**
		(0.034)	(0.049)
dist_quintile5		-0.520***	-0.420***
		(0.064)	(0.058)
share_fb_nC:dist_quintile2			0.001
			(0.001)
share_fb_nC:dist_quintile3			-0.0003
			(0.001)
share_fb_nC:dist_quintile4			-0.002*
			(0.001)
share_fb_nC:dist_quintile5			-0.003*
			(0.002)
MURs	2.000***	2.100***	2.100***
	(0.150)	(0.150)	(0.150)
Time FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Observations	46,534	46,534	46,534
Adjusted R^2	0.170	0.170	0.170

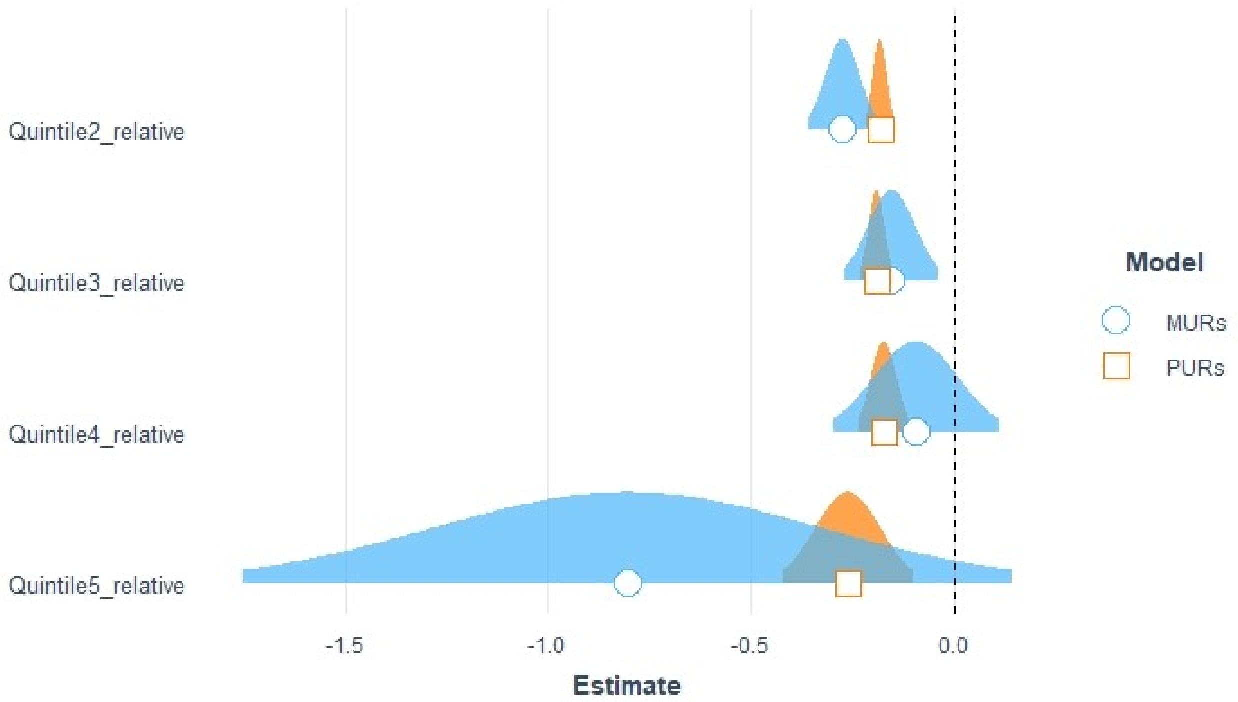
Notes: Dependent variable is the share of firm birth in the municipalities. Significant levels: ***p < 0.001; **p < 0.01; *p < 0.05.

Figure A2.1: Post-Estimation Plot of Alternative Distance Quintiles



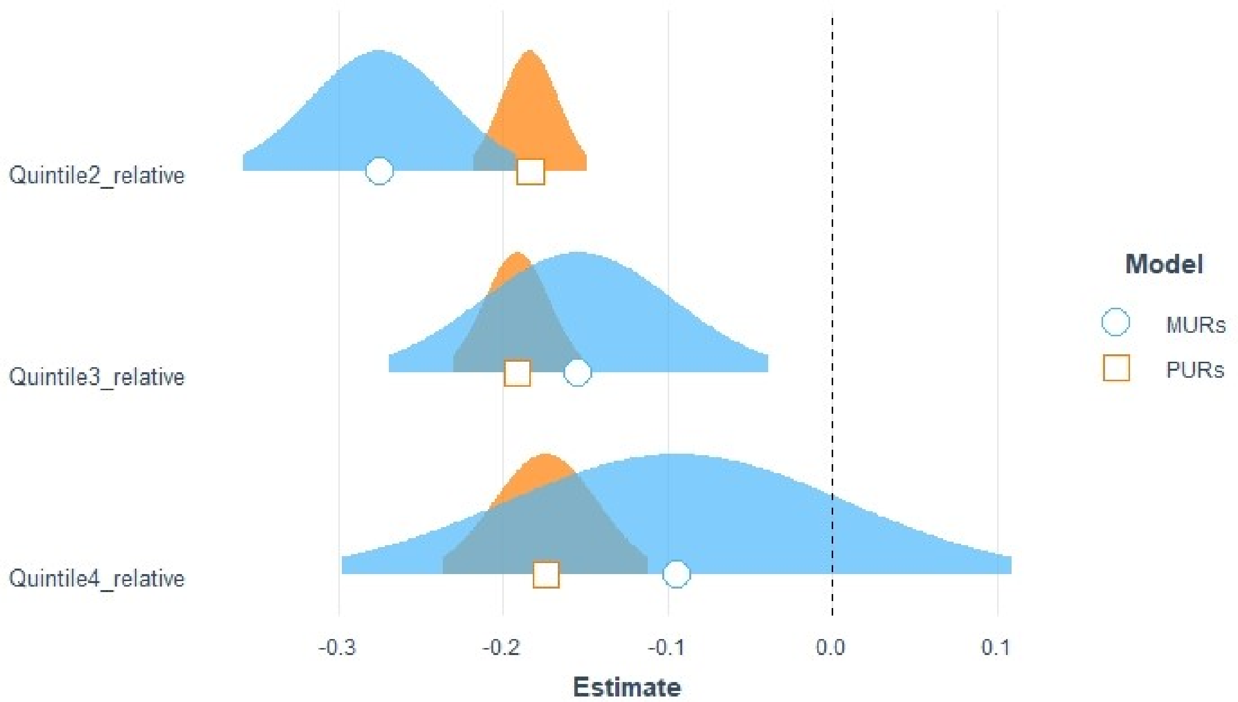
Notes: Post-estimation results from the model presented in Table A2.6 (without interactions). The filled areas display the distribution of the standard errors.

Figure A2.2: Post-Estimation Plot Distance Quintiles like Baseline



Notes: Post-estimation results from the model presented in Table A2.5 (.4 and .5) (without interactions). The filled areas display the distribution of the standard errors.

Figure A2.3: Post-Estimation Plot Distance Quintiles like Baseline Quntiles 2-4



Notes: Post-estimation plot of the distance quintiles as presented in the baseline models Quntiles 2-4

Table A2.8: Absolute Firm Birth Patterns

	Full Sample (1)	MUR (2)	PUR (3)
Population: 10,000 inhabitants	0.613***	0.627***	0.607***
	(0.020)	(0.013)	(0.026)
University Dummy	0.896***	0.841***	0.817***
	(0.096)	(0.107)	(0.157)
Population: 10,000 inhabitants next Center	0.003***	-0.020***	0.0002
	(0.0005)	(0.005)	(0.001)
Universities next Center	-0.015**	0.054***	-0.053***
	(0.006)	(0.010)	(0.008)
Absolute firm birth Region	0.001***	0.001***	0.0004***
	(0.0001)	(0.0001)	(0.0001)
Share of firm birth next Center	-0.002***	-0.007***	0.032***
	(0.001)	(0.001)	(0.003)
Distance to next center	-0.031***		
	(0.001)		
dist quintile2		-0.479***	-0.387***
		(0.053)	(0.060)
dist quintile3		-0.445***	-0.372***
		(0.074)	(0.059)
dist quintile4		-0.732***	-0.550***
		(0.207)	(0.071)
dist quintile5		-0.409	-1.020***
		(0.787)	(0.117)
share fb nC:dist quintile2		0.001	-0.015***
		(0.001)	(0.002)
share fb nC:dist quintile3		-0.003***	-0.016***
		(0.001)	(0.002)
share fb nC:dist quintile4		-0.004	-0.017***
		(0.004)	(0.002)
sharefb nC:dist quintile5		-0.008	-0.011***
		(0.015)	(0.003)
Time FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Observations	46,534	31,722	14,812
Adjusted R^2	0.442	0.379	0.480

Notes: Dependent variable is the absolute number of firm birth. Column (1) refers to the baseline with the full dataset, Column (2) shows results for the subset of monocentric regions and Column (3) shows the subset of polycentric regions. Cluster Robust Standard Errors in parenthesis on municipality-level. Significant levels: ***p < 0.001; **p < 0.01; *p < 0.05.

Figure A2.4: Post-Estimation Interaction Plot Absolute Quintiles in MURs

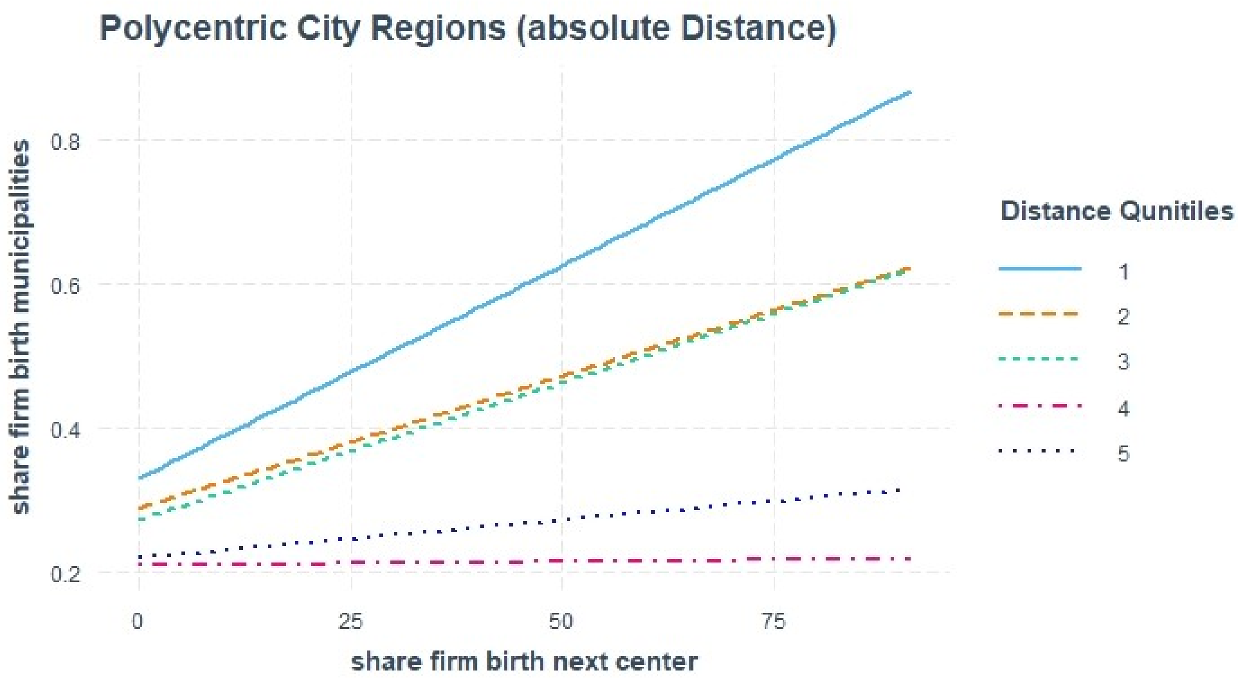
Notes: The figure shows the estimation results of the interaction terms similar to the baseline model presented in Table A2.6 with the alternative quintile definition for MURs.

Table A2.9: Robustness Check with Municipality Characteristics

	FullSample (1)	MUR (2)	PUR (3)
Population: 10,000 inhabitants	0.378***	0.881***	0.207***
	(0.013)	(0.026)	(0.008)
Absolute firm birth Region	-0.0002***	-0.0004***	-0.0001**
	(0.0001)	(0.0001)	(0.00004)
University Dummy	0.887***	0.881***	0.316***
	(0.111)	(0.188)	(0.070)
Distance to next center	-0.007***	-0.009***	-0.007***
	(0.001)	(0.002)	(0.0005)
Time FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Observations	46,534	31,722	14,812
Adjusted R^2	0.158	0.167	0.378

Notes: Dependent variable is the share of firm birth in the municipalities. Significant levels: ***p $<$ 0.001; **p $<$ 0.01; *p $<$ 0.05.

Figure A2.5: Post-Estimation Interaction Plot Absolute Quintiles in PURs



Notes: The figure shows the estimation results of the interaction terms similar to the baseline model presented in Table A2.6 with the alternative quintile definition for PURs.

A3 Appendix 3

Table A3.1: Summary Statistics for Location Characteristics

Statistic	N	Mean	St. Dev.	Min	Max
year	16,520	2,012.50	2.29	2,009	2,016
households	16,520	1,838.19	2,358.65	0	15,539
commercial buildings	16,520	285.82	514.45	0	9,708
light rail	16,520	0.98	3.29	0	32
bus stops	16,520	9.63	9.68	0	68
motorway crossing	16,520	0.18	0.77	0	10
amenities	16,520	11.03	34.49	0	438
prices	14,544	505.86	131.56	225.36	1,490.69
distance	16,520	11,414.61	6,054.14	39.07	107,234.60

Notes: Summary Statistics are pooled over the sample period from 2009 to 2016.

Table A3.2: Model (1) with Control Variables and Alternative Dependent Variables

	ln(firmbirth)	Share firmbirth	ln(in-mover)	Share in-mover	ln(firmbirth) only positive
	(1.1)	(1.2)	(1.3)	(1.4)	(1.5)
N ln(STOCK_LAG)	0.289***	0.145***	0.230***	0.160***	0.212***
	(0.007)	(0.007)	(0.007)	(0.009)	(0.010)
N* ln(STOCK_LAG)	0.036***	-0.005*	0.030***	0.002	0.012
	(0.004)	(0.003)	(0.003)	(0.004)	(0.009)
N Research Inst.	0.029**	0.057**	0.001	0.014	0.026**
	(0.014)	(0.025)	(0.017)	(0.030)	(0.013)
N* Research Inst	0.012**	0.040***	-0.001	0.028***	0.006
	(0.005)	(0.007)	(0.005)	(0.008)	(0.005)
N Higher Educaion Inst.	0.104***	0.077***	0.119***	0.147***	0.072***
	(0.014)	(0.019)	(0.017)	(0.030)	(0.012)
N* Higher Educaion Inst.	0.048***	0.060***	0.063***	0.079***	0.042***
	(0.006)	(0.007)	(0.006)	(0.010)	(0.006)
ln(commercialbuildings)	0.049***	0.046***	0.032***	0.060***	0.133***
	(0.006)	(0.008)	(0.005)	(0.010)	(0.014)
ln(househoulds)	-0.026***	-0.034***	-0.042***	-0.055***	-0.061***
	(0.004)	(0.006)	(0.003)	(0.007)	(0.009)
Light-rail	0.002	-0.001	-0.0001	-0.002	0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Busstations	-0.0004	-0.007***	-0.002***	-0.007***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Motorways	0.011**	-0.003	0.008	-0.003	0.009
	(0.005)	(0.002)	(0.005)	(0.003)	(0.006)
Amenities	0.005***	0.004***	0.004***	0.004***	0.004***
	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0002)
ln(prices)	0.163***	-0.142***	0.091***	-0.207***	-0.019
	(0.030)	(0.040)	(0.027)	(0.055)	(0.031)
Distance to CBD	-0.00000**	0.00000	0.00000	0.00000	-0.00001***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
City FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	12,726	12,726	12,726	12,726	4,580
Adjusted R^2	0.700	0.598	0.640	0.514	0.717

Notes: Cluster Robust Standard Errors in parenthesis on grid-level. Significant levels: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Table A3.3: Model (2) Full Tables and Alternative Dependent Variables

	ln (firmbirth) (2.1)	Share firmbirth (2.2)	ln (in-mover) (2.3)	Share in-mover (2.4)	ln(firmbirth) only positive (2.5)
N Research Inst. - IT	0.079	-0.071*	-0.024	0.026	-0.223***
	(0.058)	(0.041)	(0.058)	(0.041)	(0.032)
N HEI - IT	0.206***	0.343***	0.022	0.280**	0.149***
	(0.059)	(0.129)	(0.059)	(0.129)	(0.052)
	(0.026)	(0.040)	(0.026)	(0.040)	(0.023)
N HEI Econ/Social	-0.006	-0.066**	0.052**	0.012	-0.028
	(0.023)	(0.030)	(0.023)	(0.030)	(0.020)
N HEI Music/Arts/Des.	0.167***	0.026	0.195***	0.044	0.153***
	(0.037)	(0.054)	(0.037)	(0.054)	(0.034)
N Research Inst. Interdiscip.	-0.125	0.504**	0.170	1.136***	0.049
	(0.135)	(0.223)	(0.135)	(0.223)	(0.128)
N Research Inst. – Med./Law	0.074	0.050	-0.001	-0.019	0.077*
	(0.048)	(0.043)	(0.048)	(0.043)	(0.045)
N HEI Med./Law	0.067	0.061	0.065	0.089	0.033
	(0.048)	(0.063)	(0.048)	(0.063)	(0.045)
N Research Inst. - STEM	-0.015	-0.075**	-0.005	-0.080**	-0.023
	(0.038)	(0.037)	(0.038)	(0.037)	(0.037)
N HEI STEM	-0.099	-0.173**	-0.168***	-0.342***	-0.074
	(0.064)	(0.084)	(0.064)	(0.084)	(0.051)
N* Research Inst. - IT	-0.026	-0.114***	-0.044	-0.090***	-0.067**
	(0.038)	(0.022)	(0.038)	(0.022)	(0.030)
N* HEI - IT	0.026	-0.048	0.029	0.003	0.031
	(0.021)	(0.030)	(0.021)	(0.030)	(0.021)
N* Research Inst. Econ/Social	0.024***	0.021	0.013	0.011	-0.008
	(0.009)	(0.019)	(0.009)	(0.019)	(0.009)
N* HEI Econ/Social	0.021***	0.024***	0.027***	0.018**	0.012
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
N* HEI Music/Arts/Des.	0.080***	-0.008	0.085***	0.018	0.046***
	(0.015)	(0.020)	(0.015)	(0.020)	(0.014)
N* Research Inst. Interdiscip.	-0.027	0.770***	0.085	0.898***	0.186***
	(0.070)	(0.193)	(0.070)	(0.193)	(0.063)
N* Research Inst. – Med./Law	0.067***	-0.005	-0.010	-0.059***	0.030*
	(0.017)	(0.018)	(0.017)	(0.018)	(0.017)
N* HEI Med./Law	0.091***	0.173***	0.096***	0.227***	0.059***
	(0.018)	(0.034)	(0.018)	(0.034)	(0.017)
N* Research Inst.- STEM	-0.006	-0.036***	-0.007	-0.046***	-0.009
	(0.011)	(0.009)	(0.011)	(0.009)	(0.011)
N* HEI STEM	-0.093***	0.052	-0.085***	0.015	-0.049**
	(0.023)	(0.032)	(0.023)	(0.032)	(0.022)
Neighborhood controls	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	12,726	12,726	12,726	12,726	4,580
Adjusted R^2	0.703	0.638	0.644	0.552	0.729

Notes: Cluster Robust Standard Errors in parenthesis on grid-level. Significant levels: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Table A3.4: Model (3) Alternative Dependent Variables

	ln(firmbirth)	Share firmbirth	ln (in-mover)	Share in-mover	ln(firmbirth) only positive
	(3.1)	(3.2)	(3.3)	(3.4)	(3.5)
n Research Universities	-0.029	-0.145***	0.031	-0.020	-0.049
	(0.044)	(0.043)	(0.041)	(0.064)	(0.039)
n Univ. of Applied Sciences	0.140***	0.119***	0.161***	0.234***	0.082***
	(0.024)	(0.038)	(0.027)	(0.053)	(0.022)
n Univ. Arts/Music/Design	0.165***	0.221***	0.119**	0.082	0.164***
	(0.046)	(0.078)	(0.056)	(0.108)	(0.041)
*n** Research Universities	0.036**	0.133***	0.058***	0.155***	0.103***
	(0.016)	(0.030)	(0.015)	(0.033)	(0.017)
*n** Univ. of Applied Sciences	0.049***	0.033***	0.049***	0.049***	0.016*
	(0.008)	(0.010)	(0.009)	(0.013)	(0.008)
*n** Univ. Arts/Music/Design	0.066***	0.069	0.118***	0.085	0.052**
	(0.024)	(0.046)	(0.026)	(0.057)	(0.022)
Neighborhood Controls	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Observations	12,726	12,726	12,726	12,726	4,580
Adjusted R^2	0.701	0.605	0.641	0.519	0.727

Notes: Cluster Robust Standard Errors in parenthesis on grid-level. Significant levels: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Table A3.5: Baseline Models without fixed effects

	ln(firmbirth) (1)	Share firmbirth (2)	ln(in-mover) (3)	Share in-mover (4)
n ln(STOCK_LAG)	0.275***	0.136***	0.230***	0.151***
	(0.008)	(0.007)	(0.007)	(0.009)
n Research Inst.	0.032**	0.057**	0.0004	0.015
	(0.014)	(0.025)	(0.017)	(0.031)
n Higher Educaion Inst.	0.105***	0.080***	0.120***	0.150***
	(0.015)	(0.020)	(0.017)	(0.030)
n ln(commercialbuildings)	0.077***	0.067***	0.032***	0.081***
	(0.006)	(0.007)	(0.005)	(0.009)
n ln(househoulds)	-0.036***	-0.036***	-0.043***	-0.056***
	(0.004)	(0.005)	(0.003)	(0.007)
n Light-rail	0.001	-0.004**	-0.0002	-0.004**
	(0.001)	(0.001)	(0.001)	(0.002)
n Busstations	-0.002***	-0.009***	-0.002***	-0.009***
	(0.001)	(0.001)	(0.001)	(0.001)
n Motorways	0.014**	-0.001	0.007	-0.001
	(0.006)	(0.003)	(0.005)	(0.003)
n Amenities	0.005***	0.004***	0.004***	0.004***
	(0.0002)	(0.0003)	(0.0002)	(0.0003)
n ln(prices)	0.076***	0.017	0.053***	-0.032
	(0.020)	(0.022)	(0.017)	(0.029)
n Distance to CBD	-0.00000***	-0.00000***	0.00000	-0.00000***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
n* ln(STOCK_LAG)	0.018***	-0.026***	0.030***	-0.019***
	(0.004)	(0.003)	(0.003)	(0.003)
n* Research Inst	0.017***	0.038***	-0.001	0.027***
	(0.005)	(0.007)	(0.005)	(0.008)
n* Higher Educaion Inst.	0.051***	0.065***	0.063***	0.084***
	(0.006)	(0.007)	(0.006)	(0.010)
City FE	No	No	No	No
Time FE	No	No	No	No
Observations	12,726	12,726	12,726	12,726
Adjusted R^2	0.693	0.590	0.640	0.507

Notes: Cluster Robust Standard Errors in parenthesis on grid-level. Significant levels: ***p < 0.001; **p < 0.01; *p < 0.05.

Table A3.6: Robustness Estimation including Campus-Dummy

	ln(firmbirth) (1)	Share firmbirth (2)	ln(in-mover) (3)	Share in-mover (4)
n Research Institution	0.042***	0.070***	0.010	0.035
	(0.015)	(0.024)	(0.017)	(0.029)
n Research Universities	-0.023	-0.124**	0.112**	0.064
	(0.048)	(0.051)	(0.048)	(0.069)
n Univ. of Applied Sciences	0.141***	0.120***	0.172***	0.243***
	(0.025)	(0.039)	(0.027)	(0.054)
n Univ. Arts/Music/Design	0.165***	0.223***	0.118**	0.084
	(0.047)	(0.079)	(0.057)	(0.110)
n Campus	-0.019	-0.067	-0.221***	-0.241**
	(0.063)	(0.082)	(0.070)	(0.112)
n^* Research Institution	0.015***	0.034***	0.002	0.028***
	(0.006)	(0.007)	(0.006)	(0.008)
n^* Research Universities	0.045**	0.168***	0.081***	0.218***
	(0.019)	(0.045)	(0.019)	(0.051)
n^* Univ. of Applied Sciences	0.050***	0.037***	0.051***	0.056***
	(0.008)	(0.011)	(0.009)	(0.014)
n^* Univ. Arts/Music/Design	0.065***	0.064	0.117***	0.078
	(0.024)	(0.047)	(0.026)	(0.058)
n^* Campus	-0.023	-0.087*	-0.062**	-0.157***
	(0.026)	(0.050)	(0.025)	(0.060)
Neighbourhood Controls	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	12,726	12,726	12,726	12,726
Adjusted R^2	0.701	0.607	0.642	0.523

Notes: Cluster Robust Standard Errors in parenthesis on grid-level. Significant levels: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

Table A3.7: Baseline Models - subsample ICT only

	ln(firmbirth) (1)	Share firmbirth (2)	ln(in-mover) (3)	Share in-mover (4)
n ln(STOCK_LAG)	0.238***	0.190***	0.199***	0.226***
	(0.008)	(0.011)	(0.007)	(0.015)
n Research Inst.	0.008	0.024	0.018	0.033
	(0.016)	(0.029)	(0.016)	(0.037)
n Higher Educaion Inst.	0.078***	0.082***	0.106***	0.183***
	(0.016)	(0.024)	(0.019)	(0.037)
n* ln(STOCK_LAG)	0.032***	-0.002	0.021***	-0.004
	(0.004)	(0.004)	(0.003)	(0.005)
n* Research Inst.	0.012**	0.032***	-0.001	0.027***
	(0.005)	(0.008)	(0.005)	(0.010)
n* Higher Educaion Inst.	0.034***	0.056***	0.039***	0.068***
	(0.006)	(0.008)	(0.006)	(0.011)
Neighbourhood Controls	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	12,726	12,726	12,726	12,726
Adjusted R^2	0.644	0.533	0.578	0.431

Notes: The sub-sample only contains firms that are formally registered in the ICT-sector as described in the Data Section 5.3. Cluster Robust Standard Errors in parenthesis on grid-level. Significant levels: ***p < 0.001; **p < 0.01; *p < 0.05.