Spatial sorting of innovative firms and heterogeneous effects of agglomeration on innovation in Germany*

Annekatrin Niebuhr[†‡] Jan Cornelius Peters[§†] Alex Schmidke[†]

[†]IAB Nord, Regional Research Network of the Institute for Employment Research, Projensdorfer Str. 82, 24106 Kiel, Germany
[‡]Empirical Labor Economics and Spatial Econometrics, Department of Economics, Christian-Albrechts-Universität zu Kiel, Olshausenstr. 40, 24098 Kiel, Germany, Annekatrin.Niebuhr@iab.de
[§]Thünen Institute of Rural Studies, Bundesallee 64, 38116 Braunschweig, Germany, Cornelius.Peters@thuenen.de

revised version: May 10, 2019

Paper prepared for
'59th ERSA Congress',
August 27-30, 2019, Lyon

Abstract

We examine the effects of urbanization and localization on four distinct types of innovation in manufacturing and services. Furthermore, estimating multilevel panel regression models, we investigate the sorting of highly innovative firms into dense urban regions by considering both observable and unobservable firm characteristics. The results indicate that spatial sorting is important. A large portion of the regional differences in firm innovation rates is due to firm characteristics. Estimates that ignore unobserved heterogeneity at the firm level still point to a positive and significant impact of localization on different types of innovation. However, once we include firm fixed effects and distinguish between manufacturing and services, only some weak indication for positive effects of localization on radical innovations of manufacturing firms remains. In addition, the rate to adopt an existing product by an manufacturing firm is positively related to the regional employment density. For the service sector, in contrast, we find adverse effects of localization on different kinds of innovation and no important effect of urbanization.

JEL classification: D22, O31, R12
Keywords: Agglomeration economies, firm innovation, localization, urbanization

*former title: "Heterogeneous effects of agglomeration on firm innovation in Germany"

Introduction

The concentration of research and development (R&D) and innovation in space is a well-established finding of regional scientists (Audretsch and Feldman 2004). There is comprehensive evidence for a positive relationship between density and regional innovation performance (e.g., Sedgley and Elmslie 2004; Andersson et al. 2005; Carlino et al. 2007). Agglomeration economies and in particular localized knowledge spillovers are supposed to give rise to the spatial clustering of innovation (Audretsch and Feldman 1996, Bottazzi and Peri 2003). Moreover, there is evidence that firms located in dense urban areas tend to show higher innovation rates than firms located in rural regions (Naz et al. 2015). These differences might be due to composition effects, i.e., innovative firms being overrepresented in large urban areas and/or effects of an urban environment, and in particular of agglomeration economies, on the innovative performance of firms.

In this paper, we aim at disentangling the impact of the two channels in order to provide new evidence on the relative importance of composition versus agglomeration effects for (regional disparities in) innovation output. We investigate whether higher firm innovation rates in urban environments can primarily be traced back to a sorting of more innovative firms into dense urban areas or whether agglomeration economies, more precisely, urbanization and localization foster the likelihood of firms to innovate. Moreover, we address that the significance of different types of agglomeration economies, urbanization and localization, presumably varies across types of innovation and sectors.

Multilevel analyses that combine firm-level information with aggregate data allow to disentangle the effects of spatial sorting, i.e. firm-level determinants of innovation, from the influence of regional factors (e.g. Beugelsdijk 2007, Naz et al. 2015). It is essential to appropriately control for firm characteristics to identify an unbiased effect of agglomeration on firm innovation. However, often studies cannot address important econometric issues such as unobserved heterogeneity at the firm and the regional level because they have to rely on a single wave of an innovation survey (e.g. Srholec 2010). Thus, robust empirical evidence on the significance of agglomeration economies for innovation performance of firms, especially in comparison to the impact of spatial sorting, is rare. Whether regional disparities in innovation outcomes are primarily due to the sorting of innovative firms across areas or due to important agglomeration economies influencing the innovation output of firms is still largely unexplored. Moreover, little is known about heterogeneous effects of agglomeration. While several studies consider differences between product and process innovation, only a few scholars distinguish different types of product innovation (e.g., Tödtling et al. 2009, Knoben 2009). Furthermore, evidence on the determinants of innovation in the service sector is scarce as most innovation surveys provide information on manufacturing firms only.

This study accounts for several limitations of previous research and contributes to the literature on the relationship between agglomeration and firm-level innovation in several ways. In contrast to most previous studies that tend to account for sorting based on observable firm characteristics, we make use of a panel data set and fixed effects models to additionally control for unobservable characteristics. Moreover, we allow for heterogeneous effects of agglomeration on innovation with respect to three different dimensions. First, we distinguish three types of product/service innovation (radical innovation, imitation, improvement) and also consider process innovation. So far, analyses tend to neglect that agglomeration presumably affects different types of innovation differently or they account for heterogeneity between product and process innovation only. Second, our analysis is not restricted to manufacturing firms, but also examines innovation in the service sector. This is important as the significance of agglomeration economies for innovative activity might differ between manufacturing and services as well. Finally, we distinguish urbanization from localization economies whereas earlier studies often investigate the overall influence of agglomeration and do not disentangle these two types of agglomeration effects (e.g. Srholec 2010).

The remainder of the paper is structured as follows. The next section reviews the literature on the relationship between agglomeration and innovation. Section 3 describes our firm level panel data set and section 4 discusses the estimation approach and econometric issues. In section 5 the results of the regression analysis are presented. The discussion and concluding remarks follow in section 6.

Literature

Theoretical arguments

An excellent discussion of different theoretical arguments related to the effects of agglomeration on innovation is provided by Baptista and Swann (1998). They differentiate between demand and supply side effects. With respect to the demand side they argue, e.g., that customers are a good source of ideas for innovation and, therefore, locating in large markets with many customers likely enhances innovation output of firms. On the supply side, labor market pooling and access to various intermediate inputs are often mentioned as strategic advantages of dense metropolitan areas.

Focusing on the underlying mechanisms of agglomeration economies, Duranton and Puga (2004) distinguish three basic channels that might give rise to positive effects of density on firm performance: sharing, matching, and learning. Carlino and Kerr (2015) discuss in more detail how agglomeration might influence innovation via these different channels. With respect to R&D activity and innovation learning, i.e. knowledge spillovers, is probably the most important mechanism. Jacobs (1969) argues that the urban environment promotes the generation and diffusion of ideas and innovations. The spatial concentration of people and firms in urban regions facilitates in particular the transfer of tacit knowledge via face-to-face interaction. Tacit knowledge is non-codifiable and cannot be formalized and written down. It is best transmitted via frequent and repeated face-to-face interaction (see Carlino and Kerr, 2015). Furthermore, Duranton and Puga (2001) argue that a diversified urban environment supports firms in their search for an ideal production process and promotes the development of new products by adopting processes from different activities.

Matching advantages might also matter for R&D activity. It might be, for instance, easier for firms which are located in thick urban labor markets to recruit workers who possess specific skills and knowledge that are essential for innovation. Another potential matching advantage referring to knowledge spillovers is discussed by Berliant et al. (2006). In their model, individuals who possess distinct types of knowledge search for partners in order to exchange ideas and create new knowledge.

Finally, the sharing mechanism includes, among others, advantages resulting from a greater variety of inputs and the common use of local public goods and facilities (Combes and Gobillon 2015). When innovative activity clusters locally, the existence of economies of scale in the production of specialized inputs and services such as patent attorneys and commercial R&D labs allows firms to reduce R&D costs and provides an environment that enables firms to quickly implement innovations (Carlino and Kerr, 2015).

However, besides different benefits there might also be adverse effect of agglomeration, i.e., being located in a dense urban area may deter firm from innovating (Combes and Duranton 2006; Gerlach et al. 2009). Due to a high competition for specialized workers in urban labor markets, some firms might struggle to fill vacancies with appropriate candidates, hence, lacking necessary skills and knowledge to successfully innovate. Baptista and Swann (1998) also discuss possible limits to agglomeration advantages caused by congestion effects and competition in input and output markets.

An important distinction refers to localization and urbanization economies. Localization effects base upon the clustering of firms belonging to a particular industry, while urbanization economies refer to advantages derived from the spatial concentration of economic activity per se. According to Marshall (1920), the spatial concentration of firms belonging to the same industry gives rise to localization economies that consist in knowledge and information spillovers, labor pooling due to thick markets for specialized skills and various backward and forward linkages (see also Andersson and Lööf 2012). However, Duranton and Puga (2001) argue that specialized places are better to conduct mass production of fully developed products than to develop new products. In contrast, urbanization economies might play an important role for the latter as discussed by Jacobs (1969).

Some authors also hypothesize heterogeneous effects of agglomeration on innovation (Knoben 2009, Tödtling et al. 2009). Tödtling et al. (2009) argue that different kinds of innovation might rely on specific knowledge sources and links. For instance, radical innovation might require first of all internal R&D resources and, as regards external knowledge, firms may draw more on scientific knowledge, generated in universities and research institutes. These knowledge sources might primarily locate in large urban areas but not necessarily in regions showing a clustering of specific industries. However, if it is primarily the recruitment of highly specialized scientists and technicians that matters for radical innovation, benefits of labor market pooling likely prevail which are available in regions characterized by a clustering of firms belonging to the same industry. Ebert et al. (2018) note that localization effects are expected to lead to incremental innovations and process innovations. These types of externalities are supposed to spur price competition and likely strengthen well-established industries at their locations.

In contrast, internal R&D resources and a highly specialized scientific staff might be less important for imitation and incremental innovation as compared to the monitoring of competitors (Malmberg and Maskell 2002) which might be easier in an industrial cluster. Tödtling et al. (2009) note that imitating or improving an existing product likely occurs in interaction with partners from the business sector. von Hippel (2007) stresses the importance of customers for bringing forward new solutions. In the manufacturing sector these spillovers might not necessarily be highly localized because these firms tend to serve a supra-regional market and often even operate in international markets. Tödtling et al. (2009) underline that innovation systems and corresponding networks are often not confined to specific areas and may have an international or even global reach. A location in a large local market might thus be less important for manufacturing as compared to service sector firms which primarily supply non-tradables in local markets. If service sector firms are generally more exposed to the impact of the regional context they might also particularly suffer to adverse effects arising from a large number of local competitors.

Thus, there are numerous urbanization and localization effects on innovation which operate via various channels and their importance likely differs across distinct types of innovation and sectors because the corresponding prerequisites for innovation success vary. In particular, the strength of learning effects may differ depending, among other things, on the source of knowledge, its main locations, and the mechanism of transmission.

We might observe a significant correlation between agglomeration and firm innovation rates since the urban environment influences the performance of all firms and/or because innovative firms benefit from a dense urban context and choose to locate primarily in these regions. Corresponding composition effects due to a spatial sorting of establishments on observable and unobservable characteristics may be based on different mechanisms.

First, existing firms might relocate to another region or start-ups might choose a location which provides a favorable environment for their business. In fact, Baptista and Swann (1998) emphasize that innovative activity is closely associated with firm entry. As regards the location choice of newly founded firms, for example, Moeller (2018) argues that local consumption amenities attracting highly skilled workers affect the location choice of creative service firms such as internet start-ups, resulting in a spatial concentration of these firms in specific areas. In contrast, the location choice of non-creative service firms seems not to be affected by corresponding features of an urban environment. More generally, Duranton and Puga (2001) suggest that firms choose a diversified urban environment to develop new products. However, for mass production of fully developed products they relocate to specialized areas where production costs are typically lower. It is important to note that such a spatial sorting on tasks does not require that entire firms switch location. Corresponding sorting may also be obtained by the devision of labor within firms with multiple sites if work units with certain functions are located at certain places. For example, the work unit engaged in R&D might be located in a metropolitan area, whereas the production sites are located in less dense regions. [1]

Second, the spatial sorting of innovative firms might also be caused by effects of an urban environment on firm survival. In particular, there is a growing literature which suggests that the effect of the regional context on survival might be moderated by firm characteristics (e.g., Pe'er and Keil, 2013; Pe'er et al., 2016; Howell et al., 2018). Findings indicate that corresponding externalities result in spatial sorting since newly founded firms benefit differently from specific features of the regional environment as shown by Renski (2011) for the U.S. Most recently, Ebert et al. (2018) provide empirical evidence for Germany. They find, inter alia, that localization has a positive effect on the survival of non-high-tech start-ups and less innovative firms, but no or even adverse effects on the survival of highly innovative start-ups. This might reduce average innovation rates in specialized regions. Knoben (2009) notes that there might be a negative selection effect in dense urban areas. For instance, Alcácer (2006) concludes that more capable firms are more likely to stop competitors from entering their market area and can more effectively force weaker firms out. Due to these effects the stronger firms end up being located in relative isolation, whereas the weaker firms end up in agglomerations since they neither drive rivals out of their local market nor deter new start-ups from locating in the same region.

Empirical evidence

As regards the voluminous empirical literature on the link between agglomeration and innovation, we refrain from a detailed review and refer to a recent survey provided by Carlino and Kerr (2015).[2] Many studies operate at the aggregate level and investigate the relationship between density or size of regions and their innovation output (e.g., Sedgley and Elmslie 2004; Andersson et al. 2005; Carlino et al. 2007). The majority of studies that investigate the issue on the aggregate regional level tends to detect a positive relationship between the size and density of regions and their innovative output. However, these studies cannot disentangle the impact of agglomeration on innovation from composition effects.

[1] Disparities in R&D activity and innovation rates between East and West Germany serve as an illustrative example in this context. As for instance discussed by Niebuhr (2017), the East-West gap in R&D activity is (partially) explained by the fact that the headquarters and R&D units of West German firms remained in West Germany after reunification and only parts of the production moved to East Germany. In the data we use, the IAB Establishment Panel, different sites of one firm are considered independent establishments.

[2] In particular, many firm level and multilevel studies investigate for selected industries specific mechanisms that might be behind the link between agglomeration and innovation (e.g., Fornahl et al. 2011 focusing on R&D subsidies, collaboration networks and biotech firms and Ben Abdesslem and Chiappini 2019 analyzing the impact of a cluster policy on productivity, employment, and total fixed assets in the French optic/photonic industry). While these analyses provide very detailed and robust evidence on specific industries and selected channels through which agglomeration might impact firm innovation, often their results cannot be generalized and it is difficult to compare their findings with studies that apply a more aggregate perspective on agglomeration economies.

An increasing number of multilevel analyses address this problem since they control for firm characteristics when estimating the effect of agglomeration on innovation (Naz et al. 2015). Using micro data from the Community Innovation Survey in the Czech Republic, Srholec (2010) shows that the quality of the regional innovation system has a significant positive effect on firm innovation. Findings by De Beule and Van Beveren (2012) suggest that agglomeration economies increase the innovation output of Belgian firms. Their estimates point to a beneficial impact of both localization and urbanization economies. However, localization economies seem to matter only for low-tech manufacturing and the service sector, while urbanization externalities enhance innovation output in high-tech service sectors. The findings indicate that it is important to distinguish between both sectors as well as urbanization and localization effects. These results are confirmed by evidence provided by Beugelsdijk (2007) and Smit et al. (2013) who detect significant effects of the regional context only for specific sectors and types of innovation. Smit et al. (2013) also stress that estimates differ considerably depending on whether firm characteristics are included or not. Thus, studies not controlling for firm characteristics suffer from severe omitted variable bias.

Many multilevel studies base on cross sectional data sets that combine firm-level information with data on regional characteristics for a specific year (e.g., De Beule and Van Beveren 2012, Smit et al. 2013). These analyses account for observed firm characteristics and, partly, unobserved heterogeneity at the regional level. But they cannot control for the spatial sorting of firms based on unobserved firm characteristics. Only a few studies control for unobserved firm characteristics via estimating fixed effects model. In this respect, robust empirical evidence is still scarce. Baptista and Swann (1998) make use of a multilevel panel data set to investigate the relationship between agglomeration and firm innovation for manufacturing firms located in the UK. They also distinguish between urbanization and localization effects and find that a firm is more likely to innovate in a region where the own industry is present above average. In contrast, there is no evidence for important urbanization effects. The authors suggest that this result might indicate the presence of congestion effects. Andersson and Lööf (2012) desist from applying a fixed-effects estimator because in their panel data set the vast majority of the variation is between rather than within firms. With fixed effects more than 90 percent of the observations are dropped because there is no variation in innovation over time. Like Bellmann et al. (2013), they prefer a random-effects model over the fixed-effects approach because the latter is not very efficient, since it relies on within variation only.

Beaudry and Breschi (2003) investigate whether firms located in industrial clusters in the UK and Italy tend to show a higher innovation output compared to firms located outside these areas. They apply a negative binomial model that accounts for unobserved heterogeneity via including previous innovation performance to identify factors that influence the number of patent at the firm level. Their results point to positive localization effects conditional on the innovation output of the local sector. While there is some evidence for important positive localization effects if the cluster is populated by innovative firms, the results also indicate adverse localization effects if local firms belonging to the same industrial sector are not innovative.

To summarize, although evidence on agglomeration effects which is based on multilevel analysis is clearly growing, the overall number of corresponding studies is still relatively small, findings are far from being unambiguous and some studies might suffer from ignored unobserved heterogeneity at the firm level. Thus, so far it is difficult to assess the relative importance of spatial sorting as opposed to agglomeration effects on firm's innovation output. Moreover, evidence on the factors that influence innovation of service sector firms is scarce.

Data

To investigate the impact of urbanization and localization on firm innovation, we make use of a linked employer-employee data set. It covers information on 18,754 establishments in Germany and refers to the period 1999-2010. An unbalanced panel data set is generated by merging data from two different sources. It represents a uniquely rich source of firm level information. Among others, it provides detailed information on firm innovation and establishment characteristics such as size, industry, and workforce composition. Since the location of the establishments[3] is also available, we can add information on the regional context, i.e., the employment density and information on the regional industry structure to the firm level data set.

The establishment level data is taken from the Establishment Panel of the Institute for Employment Research (IAB) and the IAB Establishment History Panel (BHP). The former is an annual representative survey. It covers 1 percent of all plants (approximately 16,000) in Germany.[4] We use the waves 2001, 2004, 2007, 2008, 2009, and 2010 which include information on firm innovation. The innovation concepts used in the IAB Establishment Panel correspond with the definitions laid down in the OSLO Manual (see OECD 2005) and the concepts used in the Community Innovation Survey (CIS). The plants are asked whether they introduced a completely new product or service during the previous year[5] (see question Q1 below, referring to the 2008 wave), whether they newly adopted a product or service (question Q2), or whether they enhanced an existing product or service (question Q3). Furthermore, for 2007, 2008, 2009, and 2010 information on process innovations is available (question Q4).

> Q1: "Have you started to offer a completely new product or service in the last business year of 2007 for which a new market had to be created?"
>
> Q2: "In the last business year of 2007, did your establishment start to offer a product/service that had been available on the market before?"
>
> Q3: "In the last business year of 2007, did your establishment improve or further develop a product or service, which had already been part of your portfolio?"
>
> Q4: "Did you develop or implement procedures in the last business year of 2007 which have noticeably improved production processes or services?"

Based on this information, we generate five different dependent variables: radical innovation referring to question Q1, imitation referring to question Q2, improvement referring to question Q3, and process innovation referring to question Q4. They take the value of 1 if the answer of an establishment to the respective question is *Yes* and 0 if the question has been answered in the negative. The fifth variable we generate 'any product or service innovation' takes the value 1 if the answer to at least one of the questions Q1 to Q3 is *Yes*.[6]

[3] Different units of one firm that are located in different municipalities are considered as independent establishments. Unfortunately, it is not possible to identify whether establishments belong to the same firm. In order to improve the readability, we use the terms firms and plants in the following as synonyms for the term establishment.

[4] See Ellguth et al. (2014) for a detailed description of the IAB Establishment Panel.

[5] Until 2007, the question refers to the previous two years.

[6] If the answer to one of the questions Q1 to Q3 is missing and the other questions have been answered in the negative, the observation is excluded from the analysis of 'any innovation'. Furthermore, we do not consider an observation in the analysis of 'improvements' and 'imitations', respectively, if a firm reports the introduction of an entirely new product or service (Q1), but no imitation (Q2, 1.4% of all observations, see Table A.2), respectively, no improvement (Q3, 3.9%). Otherwise the dependent variables would take the value 0, even though the firm is highly innovative. 'Improvement' is set to missing, in addition, if a firm reports an imitation, but no improvement (4.5%).

Many studies use patent data to measure innovation. However, Carlino and Kerr (2015) note that patents only reflect the first stage of innovation, i.e. the invention, and provide no information on the commercialization of a new product or service. Moreover, Pakes and Griliches (1980) note that not all new innovations are patented, patents significantly differ in their value and economic impact and might be subject to strategic behavior. Feldman and Kogler (2010) conclude that studies which focus on invention to provide evidence on innovation output should be interpreted with caution.[7]

Smith (2006) notes that the number of studies using innovation survey data is growing rapidly. In particular, the CIS has become an important data source for innovation studies and seems to provide very reliable information for manufacturing (Smith 2006). A key concern discussed in the literature is whether the CIS approach which is also used for the IEB Establishment Panel should be applied to the service sector because it was originally developed for manufacturing. Smith (2006) argues that the technological definitions of change might be to narrow for an extremely heterogeneous services sector and its often intangible outputs. However, the questionnaire of the IAB Establishment Panel includes four questions on innovation and therefore allows to differentiate between several types of innovation. Moreover, the survey explicitly distinguishes between products and services. Therefore we are confident that applying the CIS approach should also provide reliable information on innovation in the service sector. When interpreting the results that differentiate between service and manufacturing we should keep in mind that the CIS was developed for investigating innovation of manufacturing firms and that evidence on innovation in the service sector is still scarce. In sum, the different indicators of innovative output have strengths and weaknesses. But altogether the IAB Establishment Panel provides high quality information on innovation output. The quality of the data at hand is confirmed by the differences in innovation rates across sectors and firm size categories that closely resemble the findings of other studies (see Table 3 and Table A.4).

Summary statistics in Table A.1 indicates that in almost 50 percent of the observations in our sample an innovation is reported. However, the frequency of different types of innovation varies significantly. While in only 10 percent of the observations a radical innovation is reported, the rate of improving an existing product/service is 46 percent. Moreover, often one type of innovation comes along with another type (Table A.2). For instance, firms that introduce radical innovations imitate and/or improve existing products or services at a rate of 90 percent.

The firm-level data available in the Establishment Panel is extended with information from the BHP which is based on mandatory social security notifications and therefore contains very reliable and high quality information. The BHP covers the total population of plants that employ at least one worker who is subject to the social security notification in Germany. We merge the information available in the Establishment Panel and the BHP via a plant identifier that is available in both data sets. Among others, the BHP provides for each plant information on the location, industry, firm age, number of employees and labor turnover, i.e. the sum of establishment entries and exits in period t divided by two times plant employment at the beginning of t. The BHP also includes information on the composition of the workforce with respect to different characteristics such as gender, age, level of educational attainment, occupation, and nationality.

We exclude public sector establishments and about one percent of the observations because information on innovation is missing. Moreover, we cannot use 6 percent of the remaining observations because the share of high- and low-skilled workers cannot be computed due to missing values in the data. The exclusion of these observation impacts only marginally on the composition of the data set with regard to innovation rates

[7] A detailed discussion of the pros and cons of different innovation measures is provided by a recent survey in Carlino and Kerr (2015) who focus on regional studies.

and the regional distribution of the establishments. With these restrictions, we end up with 42,412 plant-year observations, though not all observations provide information on each kind of innovation (see number of observations in Table A.1). Since only 5 percent of the establishments participated in the survey in each of the considered six years, the corresponding panel data set is unbalanced. The median number of observations per establishment amounts to 2 if we consider the full sample and 3 if we focus on establishments with at least two observations when applying fixed effects estimation.

To analyze the impact of urbanization and localization on firm innovation, we extend the firm level data set with information on the regional context focusing on employment density and the regional employment share of the industry an establishment belongs to. Corresponding regional employment figures are taken from employment statistics of the Federal Employment Agency. We distinguish 86 industries at the two-digit level and assign each establishment to one out of 141 so called labor market regions defined by Kosfeld and Werner (2012). The regions are supposed to represent integrated local labor markets. Their delineation is – like the definition of, e.g., travel to work areas (TTWA) in the UK – based on commuter flows. Corresponding regions, therefore, reflect "the regional scale of regular interactions" (Ebert et al., 2018: p.3). The labor market regions defined by Kosfeld and Werner (2012) unify NUTS3-regions connected by intense commuting. Weighted by the number of observations in our sample, their average size amounts to about 3500 sqkm which corresponds to a radius of about 32.5 km if the regions were circular (compare Table A.3).

Empirical Strategy

We begin with a comparison of unconditional average innovation rates distinguishing urban, intermediate, and rural regions for the five types of innovation defined above. Comparing these differences with the variation in average innovation rates conditional on observable firm characteristics, provides insights on the importance of composition effects for spatial disparities in innovation outcomes.

Following Naz et al. (2015), we then decompose the variation in firm innovation rates via estimating a simple model that only includes an overall intercept γ_0 and the residuals at firm level, $e_{i,r,t}$, and the regional level, μ_{0r}:

$$y_{i,r,t} = \gamma_0 + \mu_{0r} + e_{i,r,t} \tag{1}$$

The dependent variable is a binary variable that takes the value of one if establishment i that is located in region r reports in year t a specific type of innovation. If the considered type of innovation is not reported, the dependent variable is zero. The model allows us to determine how much of the variation in firm innovation rates is between regions. We use the estimates of σ_e^2 and $\sigma_{\mu_0}^2$ to calculate the intraclass correlation ρ, a measure of the dependence of firm innovation rates observed in the same region. It gives the share of the unconditional variation in innovations rates that is between regional labor markets:

$$\rho = \sigma_{\mu_0}^2 / (\sigma_{\mu_0}^2 + \sigma_e^2) \tag{2}$$

The decomposition approach gives a first indication of the significance of factors at each level for firm innovation output.

To analyze the impact of urbanization and localization economies on different types of innovation, we apply a linear probability model that contains information on the firm and the regional level. We estimate two different specifications. The first one is given by:

$$y_{i,r,j,t} = \alpha + X_{i,t-1}\beta + Z_{r,j,t-1}\delta + d_r + d_j + d_t + e_{i,r,t}. \tag{3}$$

As emphasized by Beugelsdijk (2007), the consideration of firm characteristics is essential to identify an unbiased effect of the regional environment on firm innovation. Therefore, we include a set of predetermined firm characteristics in our model. $X_{i,t-1}$ contains information on the structure of a firm's workforce, among others the employment shares of high-skilled labor and of workers engaged in R&D. In addition, it includes the size and age of the firm and labor turnover. $Z_{r,j,t-1}$ contains our pivotal explanatory variables measuring the degrees of urbanization and localization in the local labor market r, i.e., the local employment density and the percentage share of industry j in the regional economy in terms of employment. We estimate their impact conditional on region and industry fixed effects, d_r and d_j to reduce the risk of an omitted variable bias. Furthermore, we include time fixed effects d_t to control for time dependent shocks in firm innovation rates.

The second specification includes firm fixed effects α_i instead of region and industry fixed effects in order to address that also unobservable firm characteristics might impact on innovation output. The corresponding model is given by:

$$y_{i,r,j,t} = \alpha_i + X_{i,t-1}\beta + Z_{r,j,t-1}\delta + d_t + e_{i,r,t}. \tag{4}$$

By including firm fixed effects, we control for all time invariant characteristics of a firm that determine its ability to generate innovations.[8] A comparison of the results of equations (3) and (4) will indicate the necessity to control for unobserved heterogeneity at the establishment level in order to obtain unbiased results for the parameters of interest, i.e., the impact of urbanization and localization economies on different kinds of firm innovation. The differences between the regression results also enable us to assess the importance of agglomeration economies for firm innovation relative to the significance of the sorting of innovative firms into specific areas. The spatial sorting of establishments might be based on observable and unobservable firm characteristics.

Results

Table 1 summarizes mean innovation rates of establishments located in urban, intermediate and rural regions, respectively.[9] In line with expectations, firms in urban and intermediate regions report innovations on average more frequently than firms in rural regions. Considering 'any innovation' as defined in Section 'data', we observe that more than 50 percent of the establishments located in urban and intermediate regions introduced at least one type of product or service innovation, while only 43 percent of the establishments in rural regions did so. Test statistics in the right panel of Table 1 indicate that these differences in innovation rates are not only economically meaningful, but also statistically significant. Lower innovation rates in rural regions are in line with previous findings on spatial disparities in R&D activity and innovation outcome as summarized by, e.g., Audretsch and Feldman (2004). While rural areas significantly lag behind, we detect no important differences between urban and intermediate regions for this broad definition of innovation that covers the different types of product/service innovation.

[8] Including firm fixed effects implies that we also control for time invariant characteristics of the regions since the establishments in our data set do not relocate between regional labor markets.

[9] The distinction between the three types of regions refers to a classification of the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) based on population density, the size of the largest city within a region, and the population share living in cities. Summary statistics of regional employment figures by region type are given in Table A.3.

Table 1: Mean comparison of unconditional innovation rates across region types

	Average firm innovation rates by type of region in %			Mean difference, percentage points (t-statistic)		
	Urban	Inter-mediate	Rural	Urban vs. Inter-mediate	Urban vs. Rural	Inter-mediate vs. Rural
Any innovation	51.17	50.41	42.94	0.76	8.23^{***}	7.47^{***}
				(1.32)	(13.70)	(11.73)
Radical innovation	10.05	10.84	8.09	-0.78^{**}	1.97^{***}	2.75^{***}
				(−2.24)	(5.62)	(7.31)
Imitation	26.05	26.51	23.23	−0.47	2.81^{***}	3.28^{***}
				(−0.91)	(5.30)	(5.81)
Improvement	48.01	47.27	39.25	0.75	8.66^{***}	7.91^{***}
				(1.27)	(14.06)	(12.13)
Process innovation	22.52	23.87	16.36	-1.35^{**}	6.15^{***}	7.50^{***}
				(−2.19)	(10.18)	(11.71)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. *Source*: IAB Establishment Panel.

Considering distinct types of innovations, we observe significantly lower innovation rates in rural areas for all kinds of outcomes. With respect to radical innovations, the average innovation rate in urban and intermediate regions is about 25 percent higher than in rural regions.

As regards disparities between urban and intermediate regions, innovation rates in intermediate regions are significantly higher for radical innovation and process innovation whereas for the other kinds of innovation the null hypotheses of equal innovation rates cannot be rejected at conventional levels. Thus, innovation output of the firms does not seem to increase steadily with size and density of the location.

Firm innovation rates might differ across region types for two reasons. First, the regional environment in urban and intermediate regions may foster the generation of innovations, e.g., due to urbanization and localization economies. Second, innovation rates might be higher due to composition effects. For instance, higher innovation rates for radical innovations in intermediate as compared to urban regions might be due to differences in regional industry structures. Industries like chemicals, machinery and equipment which report radical innovations at rates almost twice as high as the average rate across all industries are overrepresented outside urban centers in intermediate regions (see Table A.4 and A.5). Moreover, firms with characteristics that are beneficial for the generation of innovations might self-select into agglomerations and intermediate regions.

Table 2 provides results of a variance decomposition as described by equation (1). For each type of innovation the variance at the firm level is larger than the variance at the region level. According to the intraclass correlation, between 1.4 percent (radical innovation) and 6.3 percent (process innovation) of the unconditional variation in innovation rates is between regions. Hence, the regional context does not seem to be very important for differences in innovation performance. However, it should not be neglected. Several studies show that regional characteristics influence firm innovation rates although their importance is moderate compared to firm-level factors (see, e.g., Smit et al. 2013, Sternberg and Arndt 2001).

Considering that establishments might select on observable firm characteristics in particular types of regions, Table 3 summarizes regression results of a linear probability model including dummy variables for urban and intermediate regions as well as observable establishment characteristics such as workforce composition, establishment size and age as well as industry. The results indicate that firm innovation rates are significantly

Table 2: Variance decomposition of firm innovation rates

	Any innovation	Radical innovation	Imitation	Improvement	Process innovation
Standard deviation at					
regional level	0.101	0.036	0.058	0.106	0.103
	(0.007)	(0.003)	(0.005)	(0.007)	(0.007)
firm level	0.492	0.295	0.433	0.489	0.401
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Intraclass correlation	0.040	0.014	0.017	0.045	0.063
Observations	42306	42430	40672	39770	27540

Note: Standard errors in parentheses. *Source*: IAB Establishment Panel, own calculations.

higher in urban than in rural regions (the reference) irrespective of the type of innovation, even if we control for composition effects by considering observable firm characteristics in the regression model. The same applies to intermediate regions with the exception that the positive effect for imitations is not statistically significant at conventional levels.

The results of the firm characteristics are reported in the lower panel of Table 3. They are similar to results obtained by other studies (see, e.g., Bellmann et al., 2013). The size of a firm as well as its share of high skilled labor are positively correlated with the probability to innovate. This is also true for the share of female workers. Moreover, young firms report improvements at a higher rate and process innovations at a lower rate than older firms. The share of R&D staff of an establishment is important for radical innovations. These findings also underline the quality of the information provided in the IAB Establishment Panel since they closely resemble the findings on the effects human capital, firm size, and firm age in Peters (2009). She uses firm level data from the Mannheim Innovation Panel (MIP), a data source that includes information on German manufacturing and the service sector.[10]

Table 3: Differences in firm innovation rates across region types conditional on observable establishment characteristics

	(1) Any innovation	(2) Radical innovation	(3) Imitation	(4) Improvement	(5) Process innovation
Region type					
Urban region	0.058^{***}	0.014^{**}	0.019^{**}	0.060^{***}	0.036^{***}
	(0.017)	(0.006)	(0.010)	(0.018)	(0.012)
Intermediate region	0.040^{**}	0.016^{**}	0.015	0.041^{**}	0.047^{***}
	(0.018)	(0.007)	(0.011)	(0.019)	(0.015)
Rural region			reference		
Establishment characteristics					
Share high skilled workers	0.125^{***}	0.058^{***}	0.044^{**}	0.147^{***}	0.024
	(0.023)	(0.014)	(0.020)	(0.025)	(0.019)
Share low skilled workers	-0.022	-0.017^{**}	-0.021^{*}	-0.022	-0.007
	(0.015)	(0.007)	(0.012)	(0.016)	(0.013)
Share of R&D staff	-0.056	0.083^{**}	0.017	-0.070	-0.056
	(0.053)	(0.037)	(0.044)	(0.056)	(0.049)
Mean age R&D staff	0.001^{***}	0.000^{***}	0.001^{***}	0.002^{***}	0.001^{***}

Continued on next page

[10] The MIP is also the German contribution to the CIS.

Table 3 – continued from previous page

	(1) Any innovation	(2) Radical innovation	(3) Imi-tation	(4) Improve-ment	(5) Process innovation
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share female workers	0.036***	0.031***	0.066***	0.031***	0.021**
	(0.013)	(0.007)	(0.012)	(0.013)	(0.010)
Labour turnover	−0.022***	−0.004***	−0.012***	−0.021***	−0.002
	(0.005)	(0.001)	(0.003)	(0.005)	(0.013)
Firm age, reference: >25 years					
<4 years	0.039	0.021	0.049**	0.048*	0.022
	(0.026)	(0.014)	(0.024)	(0.026)	(0.019)
4-6 years	0.032**	0.025***	0.039***	0.030*	0.032**
	(0.016)	(0.007)	(0.011)	(0.017)	(0.013)
7-10 years	0.016	0.012**	0.024**	0.016	0.011
	(0.014)	(0.006)	(0.010)	(0.014)	(0.011)
11-15 years	−0.023*	0.002	0.000	−0.024*	−0.019
	(0.012)	(0.005)	(0.009)	(0.013)	(0.012)
16-20 years	−0.042***	−0.003	−0.013	−0.042***	−0.045***
	(0.015)	(0.007)	(0.011)	(0.016)	(0.012)
21-25 years	0.011	0.007	−0.006	0.009	−0.004
	(0.013)	(0.008)	(0.011)	(0.014)	(0.013)
Firm size, reference: >500 emp.					
1-10 emp.	−0.384***	−0.150***	−0.224***	−0.407***	−0.443***
	(0.018)	(0.014)	(0.018)	(0.018)	(0.020)
11-20 emp.	−0.295***	−0.132***	−0.166***	−0.317***	−0.383***
	(0.017)	(0.014)	(0.018)	(0.017)	(0.022)
21-50 emp.	−0.226***	−0.111***	−0.134***	−0.245***	−0.348***
	(0.016)	(0.015)	(0.018)	(0.017)	(0.020)
51-100 emp.	−0.182***	−0.086***	−0.107***	−0.192***	−0.310***
	(0.016)	(0.015)	(0.018)	(0.017)	(0.021)
101-200 emp.	−0.137***	−0.086***	−0.108***	−0.142***	−0.239***
	(0.016)	(0.014)	(0.017)	(0.017)	(0.022)
201-500 emp.	−0.068***	−0.048***	−0.050***	−0.066***	−0.147***
	(0.014)	(0.014)	(0.018)	(0.014)	(0.025)
Observations	42306	42340	40672	39770	27540
Adj. R^2	0.154	0.064	0.074	0.181	0.145

Note: Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
All models include time and industry fixed effects. *Source*: own calculations,
IAB Establishment Panel, Establishment History Panel & Federal Employment Agency.

Comparing the coefficients in Table 3 with the unconditional differences in mean innovation rates reported in Table 1 confirms that a significant part of the unconditional differences is due to selection effects as emphasized by Beugelsdijk (2007). The difference in regional innovation rates are substantially lower if establishment characteristics are considered.

To analyze the significance of urbanization and localization economies for different kinds of firm innovation, we replace the dummy variables for the type of region by the regional employment density and the employment share of the industry an establishment belongs to.[11] Furthermore, we control for time invariant differences in innovation rates across regions and sectors by means of region and industry fixed effects (see equation (3)). Hence, in this specification the pivotal parameters are estimated based on the variation in innovation outcomes within establishments over time and across establishments located in the same region and belonging to the same

[11] Alternative specifications including spatial lags of employment density and industry share do not point to significant spatial spillover effects across the borders of the considered labor market regions (Table A.6). Therefore, we focus on the impact of employment density and the industry share within the region in which an establishment is located.

industry.[12] The significance of urbanization economies is identified based on changes in regional employment densities over time and the impact of localization economies based on the spatial variation in the share of an industry and corresponding changes across years.

The results summarized in Table 4 suggest that urbanization economies are not related to higher innovation rates. Though the estimated parameter of regional employment density is positive for each type of innovation, it is not statistically significant at the five percent level. However, there is weak evidence for a positive effect on imitation in column (3). In contrast, for the regional industry share, we obtain in column (1) a parameter that is statistically significant at the 1 percent level and economically meaningful in addition. Specifications (2)-(5) indicate that there is some heterogeneity between different kinds of innovation. For instance, an increase in the regional industry share by one standard deviation (3.8 percentage points) is associated with a 1.2 percentage point higher rate of imitation and an increase in the likelihood to improve an existing product or service by 2.1 percentage points. As regards radical innovations we cannot reject the null hypothesis that localization economies do not matter.

Table 4: Impact of urbanization and localization on innovation rates conditional on time variant establishment characteristics and region fixed effects

	(1) Any innovation	(2) Radical innovation	(3) Imi-tation	(4) Improve-ment	(5) Process innovation
ln(Employees per km^2)	0.203	0.028	0.160*	0.180	0.159
	(0.134)	(0.051)	(0.085)	(0.139)	(0.352)
Regional share of own industry	0.0051***	-0.0001	0.0032**	0.0056***	0.0048***
	(0.0014)	(0.0008)	(0.0013)	(0.0014)	(0.0016)
Observations	42306	42340	40672	39770	27540
Adj. R^2	0.171	0.071	0.083	0.200	0.170

Note: Standard errors clustered at regional level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models include time variant establishment characteristics (see results in Table A.7) as well as region and industry fixed effects (equation (3)). *Source*: own calculations, IAB Establishment Panel, Establishment History Panel & Federal Employment Agency.

So far, we analyzed the correlation between firm innovation rates and agglomeration conditional on observable establishment characteristics. However, unobservable factors might impact innovation rates as well. Taking this into account, we include establishment fixed effects in our regression analysis as described by equation (4). In this regression model, the identification of parameters rests solely on the establishments with varying innovation status. Hence, the analysis is restricted to establishments that notify a certain type of innovation in at least one year and report no innovation in at least one year. It implies that highly innovative establishments which report a certain kind of innovation every year are not considered any longer. The same applies to establishments that never report this type of innovation. The shares of these establishments differ across the different kinds of innovations as well as types of regions. The percentage of establishments that never (always) report a certain kind of innovation are on average highest (lowest) in rural areas and lowest (highest) in intermediate regions (Table A.8).

Table A.9 in the appendix summarizes results of a regression analysis where we estimate the same specification as before (without firm fixed effects, see equation (3)), but based on the reduced sample of establishments that shows a variation in the innovation status. Depending on the kind of innovation, the reduced sample

[12] Recall that the analyzed panel is highly unbalanced. The median number of observations per establishment is 2. Hence, at different points in time we observe different establishments in one particular region.

comprises between 17 percent (radical innovation) and one third (imitation) of the initial sample. For the reduced sample we do not detect a significant relationship between localization economies and firm innovations rates. The estimated coefficients are much smaller than those reported in Table 4 and not statistically significant at conventional levels. The comparison indicates that the important localization effects reported in Table 4 are driven by establishments that always or never report an innovation. These results suggest that the share of establishments within an industry that always (never) report an imitation, improvement or process innovation is higher (lower) in regions that are specialized on the respective industry. However, the innovation rates of establishments that sometimes report an innovation are not systematically higher in these regions.

In contrast, if we control of unobserved heterogeneity at the establishment level by means of firm fixed effects, we obtain highly significant negative localization effects for all kinds of product or service innovations (Table 5). The strongest effect is estimated for radical innovation. Accordingly, an increase in the share of the own industry by one standard deviation coincides with a 6.5 percentage points lower likelihood to report an entirely new product or service. One potential explanation is that an increasing specialization may also enhance the competition in sector-specific factor markets, as argued by Baptista and Swann (1998), such that establishments that do not belong to the group of highly innovative establishments face severe constraints with respect to R&D investment which might significantly reduce their potential to become innovative.[13]

Table 5: Impact of urbanization and localization on innovation rates conditional on time variant establishment characteristics and establishment fixed effects

	(1) Any innovation	(2) Radical innovation	(3) Imi- tation	(4) Improve- ment	(5) Process innovation
ln(Employees per km^2)	-0.004	-0.126	0.074	0.018	-0.273
	(0.136)	(0.155)	(0.114)	(0.145)	(0.222)
Regional share of own industry	-0.0061*	-0.0172***	-0.0069*	-0.0084**	0.0051
	(0.0031)	(0.0052)	(0.0036)	(0.0034)	(0.0048)
Observations	15395	7370	13355	12751	6940
Adj. R^2	0.020	0.039	0.028	0.021	0.032

Note: Standard errors clustered at regional level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models include time variant establishment characteristics reported in Table 3 and establishment fixed effects (equation (4)). Excluding establishments that always or never reported the respective type of innovation. *Source*: own calculations, IAB Establishment Panel, Establishment History Panel & Federal Employment Agency.

The results in Table 6 indicate that adverse effects of regional specialization on the own industry on innovation outcome of establishments with varying innovation status are driven by the service sector.[14] Focusing on manufacturing firms, we do not detect important negative localization effects. In fact, for radical innovations, we obtain a positive coefficient indicating that an increase in regional specialization on the own industry by one standard deviation comes along with a 7 percentage point higher likelihood to generate an entirely new product (or service). However, this effect is not very precisely estimated as it is significant at the 10 percent level only.

The results for manufacturing suggests, furthermore, that for these firms the likelihood of imitation is positively influenced by urbanization economies. The point estimate suggests that doubling the regional employment density comes along with a 46 percentage points ($0.667 \times \ln(2) \times 100\%$) higher likelihood of an imitation.

[13] The results are robust with regard to the inclusion of the regional share of employment engaged in R&D (see Table A.10).

[14] Corresponding sector-specific results of regressions including region and industry fixed effects and excluding establishment fixed effects by sector are summarized in Table A.11.

Table 6: Impact of urbanization and localization on innovation rates conditional on time variant establishment characteristics and establishment fixed effects, by sector

	(1) Any innovation	(2) Radical innovation	(3) Imi-tation	(4) Improve-ment	(5) Process innovation
			Services only		
ln(Employees per km^2)	-0.056	0.004	0.035	-0.065	-0.402**
	(0.162)	(0.200)	(0.161)	(0.167)	(0.200)
Regional share of own industry	-0.0070*	-0.0207***	-0.0077*	-0.0058	-0.0029
	(0.0038)	(0.0063)	(0.0044)	(0.0042)	(0.0057)
Observations	7914	2911	6303	6472	3007
Adj. R^2	0.022	0.035	0.029	0.021	0.019
			Manufacturing only		
ln(Employees per km^2)	0.335	0.161	0.667**	0.462	-0.529
	(0.331)	(0.322)	(0.291)	(0.355)	(0.706)
Regional share of own industry	0.0073	0.0199*	-0.0037	0.0010	0.0243
	(0.0128)	(0.0112)	(0.0122)	(0.0149)	(0.0157)
Observations	4878	3875	5212	4199	3092
Adj. R^2	0.023	0.043	0.031	0.028	0.050

Note: Standard errors clustered at regional level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models include time variant establishment characteristics reported in Table 3 and establishment fixed effects (equation (4)). Excluding establishments that always or never reported the respective type of innovation. *Source*: own calculations, IAB Establishment Panel, Establishment History Panel & Federal Employment Agency.

Discussion and concluding remarks

In this paper we analyze the impact of urbanization and localization economies on different kinds of firm innovation in Germany. In contrast to the majority of previous studies, we address the non-random sorting of establishments across space by means of firm fixed effects. The significant differences between the specifications with and without firm fixed effects suggest that evidence on the significance of agglomeration provided by cross-sectional multilevel analyses might not be robust.

First of all, our results indicate that there are significant differences in average firm innovation rates across regions. While disparities between urban and intermediate regions are moderate or non-existent, rural regions significantly lag behind regarding all considered kinds of innovation. With regard to the factors behind these differences, our results indicate that firm characteristics are more important than the regional context. This is in line with evidence provided by Beugelsdijk (2007), Sternberg and Arndt (2001) and Naz et al. (2015). Observable firm characteristics explain a significant part of the differences in average innovation rates between the three types of regions. Unobservable firm characteristics apparently add to the disparities as well. Hence, the findings suggest that a significant proportion of regional disparities in innovation output is caused by composition effects. Firms with characteristics that promote the generation of innovations are often located in dense regions. Sorting of innovation-prone firms into specific locations seems to be much more important to explain spatial patterns of innovation than agglomeration externalities that impact on firm's innovation output (see also Smit et al. 2013).

While our results challenge the findings of some previous studies which suggest important positive effects of agglomeration on firm innovation, they are perfectly in line with a main result of the urban scaling literature pointing to superlinear relationships between city size and different urban metrics (see Bettencourt et al. 2007a). As regards innovation, Bettencourt et al. (2007b) consider the relationship between patenting activity and city population. They detect a significant positive correlation between the number of patents and city size which might be caused by inventors being more productive in larger cities or an above average share of inventors in large urban areas. The authors find that larger urban areas have a disproportionate share of inventors, but agglomeration effects do not significantly affect the productivity of the individual inventor. Likewise, our findings suggest that the sorting of innovation-prone firms mainly drives the above-average innovation output of intermediate regions and agglomerations. However, while Bettencourt et al. (2007b) focus on urbanization effects, we also consider the impact localization economies.

Our regression results are also partly in line with the hypotheses on heterogeneous effect of agglomeration discussed in section 2. As regards the impact of urbanization on firm innovation, we do not find robust evidence that an increase in the regional employment density coincides with an increase in firm innovation rates. One exception is the rate to adopt an existing product by manufacturing firms for which we detect a positive relationship with regional employment density. This result confirms arguments put forth by Tödtling et al. (2009) and von Hippel (2007) who note that imitation might be fostered by interaction with customers and partners from the business sector in large urban markets. In line with results by Baptista and Swann (1998), there is some weak indication for a positive effect of localization economies on radical innovations of manufacturing firms which might be based on knowledge spillovers, labor pooling due to a highly specialized regional workforce, and/or various forward and backward linkages as discussed by, e.g., Marshall (1920) and Andersson and Lööf (2012). For the service sector, in contrast, we detect adverse effects of localization on firm innovation. One explanation might be an increase in the competition on input markets which might hinder R&D investments and the generation of innovation. This also confirms findings by Baptista and Swann (1998) who interpret corresponding results as indicating congestion effects. In contrast, Smit et al. (2013) find little support for important competition effects on innovation.

Our findings are not in line with the presumption that the local environment is generally less important for R&D activity in manufacturing as compared to services because the latter might primarily supply non-tradables in local markets whereas the former tend to operate more often beyond the borders of a local market area. In contrast, the negative localization effects that we observe for different types of product innovation in the service sector indicate that these firms might suffer more from a highly competitive local market than manufacturing firms.

One drawback of models that include firm fixed effects is that the estimates are only based on firms that sometimes report an innovation and sometime not. Therefore, the 'general' effect is more precisely estimated when both within and between variance is used (Bellmann et al., 2013). Estimating regressions with and without firm fixed effects suggest that there is a positive (negative) correlation between localization and the share of establishments that always (never) report a certain kind of innovation. Unfortunately, we are not able to analyze whether these differences in the distribution of highly innovative establishments can entirely be explained by composition effects or whether some of these firms indeed benefit from localization economies. For instance, firms that report in each period a radical innovation might benefit in a sense that they are kept innovative or that the number of innovations per year increases due to localization effects. Tavassoli and Karlsson (2018) provide first evidence on a significant correlation between agglomeration and the persistence of firm level innovation in Sweden. We leave an in-depth analysis of this relationship for the German context as an issue for future research.

In addition, we do not analyze to which extent the location choice of firms, for example highly innovative start-ups, is influenced by regional characteristics such as the degree of urbanization or localization (see, e.g., Moeller, 2018). One interpretation of the positive correlation between the share of highly innovative establishments and the employment share of the own sector is that entrepreneurs with high abilities to innovate favor regions which offer localization advantages or that there are more innovative spin-offs in highly specialized regions. For instance, Audretsch and Feldman (2004) highlight the role of entrepreneurship as a mechanism of knowledge spillovers. And Qian and Haynes (2014) note that start-ups gain importance for knowledge-based economic development.

Appendix

Table A.1: Summary statistics of firm level data

	N	Mean	SD	Min	Max
Type of innovation					
Any product/service innovation	42306	0.487	0.500	0.000	1.000
Radical innovation	42340	0.098	0.297	0.000	1.000
Imitation	40672	0.254	0.436	0.000	1.000
Improvement	39770	0.455	0.498	0.000	1.000
Process innovation	27540	0.213	0.409	0.000	1.000
Workforce composition and turnover					
Share of R&D labor	42412	0.020	0.077	0.000	1.000
Mean age of R&D labor	42412	9.588	18.280	0.000	62.000
Share of high-skilled labor	42412	0.083	0.173	0.000	1.000
Share of female labor	42412	0.408	0.335	0.000	1.000
Share of low-skilled labor	42412	0.092	0.180	0.000	1.000
Labor turnover	42412	0.169	0.440	0.000	62.667
Firm age					
$<$4 years	42412	0.010	0.099	0.000	1.000
4-6 years	42412	0.087	0.282	0.000	1.000
7-10 years	42412	0.129	0.335	0.000	1.000
11-15 years	42412	0.201	0.401	0.000	1.000
16-20 years	42412	0.208	0.406	0.000	1.000
21-25 years	42412	0.050	0.217	0.000	1.000
$>$25 years	42412	0.316	0.399	0.000	1.000
Firm size					
1-10 emp.	42412	0.383	0.486	0.000	1.000
11-20 emp.	42412	0.136	0.342	0.000	1.000
21-50 emp.	42412	0.176	0.380	0.000	1.000
51-100 emp.	42412	0.110	0.313	0.000	1.000
101-200 emp.	42412	0.076	0.264	0.000	1.000
201-500 emp.	42412	0.073	0.260	0.000	1.000
$>$500 emp.	42412	0.045	0.208	0.000	1.000

Source: own calculations, IAB Establishment Panel & Employment History Panel.

Table A.2: Innovation pattern

Type of product/service innovation			
Radical innovation	Imitation	Improvement	Share in %
Yes	Yes	Yes	5.3
Yes	Yes	No	0.6
Yes	No	Yes	3.1
No	Yes	Yes	13.9
Yes	No	No	0.8
No	Yes	No	4.5
No	No	Yes	20.2
No	No	No	51.1
At least one missing value			0.5
Total (N=42412)			100.0

Source: own calculations, IAB Establishment Panel.

Table A.3: Summary statistics of regional characteristics by type of region

	N	Mean	SD	Min	Max
Urban region					
ln(Employees per sqkm)	17763	5.0	0.6	3.9	6.5
Industry share in regional employment in %	17763	4.4	3.7	0.0	16.8
Radius* in km	11263	34.1	10.0	10.3	51.8
Intermediate region					
ln(Employees per sqkm)	13386	4.2	0.4	3.2	5.3
Industry share in regional employment in %	13386	4.1	3.8	0.0	42.5
Radius* in km	11263	31.5	7.7	14.3	42.1
Rural region					
ln(Employees per sqkm)	11263	3.4	0.5	2.4	4.4
Industry share in regional employment in %	11263	4.3	3.9	0.0	22.5
Radius* in km	11263	31.3	6.9	14.4	47.2

* Under the assumption that the regions are circular.

Note: The spatial units are the 141 labor market regions as defined by Kosfeld and Werner (2012). The regions are weighted by the number of observations (N). Each region is considered either as urban, intermediate, or rural region. The types of regions were defined by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR) on the basis of the 96 so called *Raumordnungsregionen* according to population density, the size of the largest city within a region, and the population share living in cities. The delineation of the 141 labor market regions and the types of regions does not always coincide, i. e., in few cases a labor market region comprises parts of two types of regions. We assign such a labor market region to one type of region according to the economic center of the region. *Source:* own calculations, Federal Employment Agency, BBSR.

Table A.4: Innovation rates by industry in percent

	Any innovation	Radical innovation	Imi-tation	Improve-ment	Process innovation
Agriculture	25.4	3.2	11.7	21.4	9.1
Mining & Quarrying	39.7	5.4	13.8	37.8	18.5
Food, Beverage & Tobacco	**58.4**	**13.7**	**35.9**	**55.6**	**24.7**
Textile, Clothes & Leather	**71.5**	**17.7**	**37.5**	**69.4**	**23.8**
Wood, Paper & Pulp	**53.1**	**10.3**	**25.9**	**50.4**	**25.2**
Chemicals	**67.3**	**18.2**	**36.4**	**65.6**	**35.2**
Metals	**57.1**	**12.5**	**28.1**	**55.0**	**31.3**
Machinery & Equipment	**69.3**	**18.3**	**35.8**	**67.9**	**35.1**
Manufacturing n.e.c.	**59.1**	**9.9**	**26.9**	**57.4**	**25.4**
Electricity, Gas, Water	36.9	5.0	15.9	33.7	17.6
Construction	29.4	3.6	15.3	24.6	9.5
Wholesale Trade & Repair	45.5	8.6	**29.4**	38.6	16.3
Retail Trade	41.8	8.6	**31.2**	31.6	12.4
Hotels & Restaurants	39.1	4.5	17.7	36.2	12.1
Transport & Communication	35.7	4.3	15.1	33.9	18.0
Financial Intermediation	**61.1**	**10.8**	**34.3**	**59.1**	**31.8**
Real Estate, Renting of Machinery	38.9	3.0	12.6	37.5	12.4
Computer & Related Services	**80.5**	**22.2**	**38.2**	**79.9**	**39.9**
Research & Development	47.1	8.1	22.6	45.1	19.9
Business Services	39.4	5.6	15.5	37.5	16.0
Environmental Services	37.3	4.5	18.9	34.9	15.0
Other services	36.6	6.4	20.6	31.8	9.7
Others	41.4	8.0	18.9	39.3	18.8
Total	48.7	9.8	25.4	45.5	21.3
Observations	42306	42340	40672	39770	27540

Note: Values above the respective average are printed in bold. *Source:* own calculations, IAB Establishment Panel.

Table A.5: Distribution of observations across industries and type of regions

	Distribution across industries	Distribution across type of regions (row share in %)		
	(share in %)	Agglo-merated	Urban	Rural
Agriculture	2.8	24.7	**32.3**	**43.0**
Mining & Quarrying	0.7	21.2	**42.1**	**36.7**
Food, Beverage & Tobacco	3.6	35.1	**36.3**	**28.7**
Textile, Clothes & Leather	1.1	24.3	**52.4**	23.3
Wood, Paper & Pulp	3.7	36.0	**34.4**	**29.6**
Chemicals	6.0	29.7	**37.4**	**33.0**
Metals	6.4	38.1	**36.7**	25.2
Machinery & Equipment	12.2	34.9	**37.7**	**27.4**
Manufacturing n.e.c.	1.7	34.8	30.9	**34.3**
Electricity, Gas & Water	1.2	39.1	31.3	**29.7**
Construction	10.1	40.4	29.0	**30.6**
Wholesale Trade & Repair	8.6	**49.4**	28.7	21.9
Retail Trade	8.2	**44.2**	30.0	25.8
Hotels & Restaurants	2.7	**44.1**	28.8	**27.1**
Transport & Communication	4.6	**50.0**	26.4	23.6
Financial Intermediation	2.3	**50.8**	**32.7**	16.5
Real Estate & Renting of Machinery	2.1	**43.6**	29.8	26.6
Computer & Related Services	1.2	**54.5**	23.8	21.7
Research & Development	11.0	**47.6**	27.9	24.5
Business Services	5.5	**58.0**	24.8	17.2
Environmental Services	0.5	40.3	27.6	**32.1**
Other services	1.7	**47.8**	25.5	**26.7**
Others	2.0	**49.3**	27.9	22.8
Total (N=42412)	100.0	41.9	31.6	26.6

Note: Values above the respective average are printed in bold. *Source:* own calculations, Establishment History Panel.

Table A.6: Impact of urbanization and localization on innovation rates conditional on time variant establishment characteristics and region fixed effects considering spatial spillover effects

	(1) Any innovation	(2) Radical innovation	(3) Imi-tation	(4) Improve-ment	(5) Process innovation
ln(Employees per km^2)	0.340**	0.053	0.169	0.358**	0.126
	(0.165)	(0.062)	(0.114)	(0.177)	(0.320)
Regional share of own industry	0.0063***	−0.0005	0.0032*	0.0069***	0.0050**
	(0.0017)	(0.0009)	(0.0017)	(0.0016)	(0.0020)
Spatial lag of …					
ln(Employees per km^2)	−0.211	−0.040	−0.015	−0.275	0.203
	(0.207)	(0.086)	(0.162)	(0.215)	(0.659)
Regional share of own industry	−0.0030	0.0010	−0.0000	−0.0034	−0.0008
	(0.0025)	(0.0012)	(0.0024)	(0.0024)	(0.0030)
Observations	42306	42340	40672	39770	27540
Adj. R^2	0.171	0.071	0.083	0.200	0.170

Note: Standard errors clustered at regional level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models include time variant establishment characteristics (see results in Table A.7) as well as region and industry fixed effects (equation (3)). The spatial lags have been computed using a row standardized contiguity weight matrix. *Source*: own calculations, IAB Establishment Panel, Establishment History Panel & Federal Employment Agency.

Table A.7: Results for the control variables

	(1) Any innovation	(2) Radical innovation	(3) Imi-tation	(4) Improve-ment	(5) Process innovation
Share high skilled workers	0.133***	0.062***	0.046**	0.155***	0.035**
	(0.020)	(0.014)	(0.020)	(0.022)	(0.017)
Share low skilled workers	−0.038**	−0.022***	−0.030**	−0.041**	−0.025*
	(0.016)	(0.007)	(0.012)	(0.016)	(0.013)
Share of R&D staff	−0.060	0.079**	0.021	−0.074	−0.045
	(0.054)	(0.037)	(0.046)	(0.056)	(0.049)
Mean age R&D staff	0.001***	0.001***	0.001***	0.002***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Share female workers	0.044***	0.029***	0.067***	0.041***	0.023**
	(0.013)	(0.007)	(0.012)	(0.014)	(0.009)
Labour turnover	−0.022***	−0.005***	−0.013***	−0.022***	−0.002
	(0.006)	(0.001)	(0.004)	(0.006)	(0.013)
Firm age, reference: >25 years					
<4 years	0.072***	0.028**	0.065***	0.086***	0.053***
	(0.026)	(0.014)	(0.024)	(0.026)	(0.017)
4-6 years	0.061***	0.031***	0.051***	0.062***	0.054***
	(0.015)	(0.007)	(0.012)	(0.015)	(0.012)
7-10 years	0.050***	0.019***	0.038***	0.053***	0.038***
	(0.013)	(0.006)	(0.010)	(0.013)	(0.010)
11-15 years	0.018	0.010*	0.017**	0.020*	0.012
	(0.011)	(0.005)	(0.009)	(0.011)	(0.010)
16-20 years	0.010	0.005	0.010	0.012	−0.002
	(0.014)	(0.007)	(0.010)	(0.014)	(0.011)
21-25 years	0.016	0.008	−0.002	0.015	0.003
	(0.014)	(0.008)	(0.012)	(0.014)	(0.013)
Firm size, reference: >500 emp.					
1-10 emp.	−0.368***	−0.147***	−0.218***	−0.392***	−0.428***
	(0.018)	(0.014)	(0.018)	(0.018)	(0.021)
11-20 emp.	−0.281***	−0.130***	−0.160***	−0.303***	−0.369***
	(0.018)	(0.014)	(0.018)	(0.017)	(0.024)
21-50 emp.	−0.213***	−0.109***	−0.129***	−0.234***	−0.334***
	(0.019)	(0.015)	(0.018)	(0.019)	(0.021)
51-100 emp.	−0.168***	−0.084***	−0.100***	−0.178***	−0.293***
	(0.017)	(0.015)	(0.018)	(0.017)	(0.021)
101-200 emp.	−0.123***	−0.084***	−0.103***	−0.130***	−0.225***
	(0.017)	(0.014)	(0.018)	(0.017)	(0.023)
201-500 emp.	−0.060***	−0.047***	−0.047***	−0.060***	−0.133***
	(0.014)	(0.014)	(0.017)	(0.014)	(0.026)
Observations	42306	42340	40672	39770	27540
Adj. R^2	0.171	0.071	0.083	0.200	0.170

Note: Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models include regional employment density and the regional share of the own industry (see results in Table 3) as well as year fixed effects, industry fixed effects and region fixed effects (equation (3)). *Source*: own calculations, IAB Establishment Panel, Establishment History Panel & Federal Employment Agency.

Table A.8: Share of establishments that never, sometimes and always report an innovation in %

	Any innovation	Radical innovation	Imi-tation	Improve-ment	Process innovation
	Establishments in urban regions				
Never	29.1	77.9	53.7	35.8	62.9
Sometimes	40.1	19.9	38.3	38.4	28.0
Always	30.8	2.2	8.0	25.8	9.1
Total	100.0	100.0	100.0	100.0	100.0
	Establishments in intermediate regions				
Never	27.5	75.4	51.3	33.1	59.8
Sometimes	41.3	22.3	40.5	40.6	30.1
Always	31.3	2.3	8.2	26.3	10.1
Total	100.0	100.0	100.0	100.0	100.0
	Establishments in rural regions				
Never	34.5	81.3	55.0	42.0	70.5
Sometimes	42.1	17.3	38.5	38.7	23.4
Always	23.4	1.4	6.6	19.3	6.0
Total	100.0	100.0	100.0	100.0	100.0
	All establishments				
Never	30.1	78.0	53.3	36.6	63.9
Sometimes	41.0	20.0	39.1	39.2	27.4
Always	28.9	2.0	7.7	24.2	8.6
Total	100.0	100.0	100.0	100.0	100.0
Establishments	9936	9945	9943	9940	7944

Note: Only establishments for which at least two observations are available are considered. *Source:* own calculations, IAB Establishment Panel.

Table A.9: Impact of urbanization and localization on innovation rates conditional on time variant establishment characteristics and region fixed effects, reduced sample

	(1) Any innovation	(2) Radical innovation	(3) Imi-tation	(4) Improve-ment	(5) Process innovation
ln(Employees per km^2)	0.316	0.102	0.373***	0.349	-0.030
	(0.219)	(0.259)	(0.162)	(0.255)	(1.220)
Regional share of own industry	0.0008	-0.0019	0.0003	0.0003	-0.0023
	(0.0018)	(0.0022)	(0.0020)	(0.0018)	(0.0025)
Observations	15395	7370	13355	12751	6940
Adj. R^2	0.028	0.029	0.028	0.030	0.022

Note: Standard errors clustered at regional level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models include time variant establishment characteristics reported in Table A.7 as well as region and industry fixed effects (equation (3)). Excluding establishments that always or never reported the respective type of innovation. *Source*: own calculations, IAB Establishment Panel, Establishment History Panel & Federal Employment Agency.

Table A.10: Impact of urbanization, localization and regional R&D activity on innovation rates conditional on time variant establishment characteristics and establishment fixed effects

	(1) Any innovation	(2) Radical innovation	(3) Imi- tation	(4) Improve- ment	(5) Process innovation
ln(Employees per km^2)	0.001	-0.114	0.149	0.014	-0.228
	(0.144)	(0.174)	(0.125)	(0.153)	(0.257)
Regional share of own industry	-0.0061*	-0.0172***	-0.0067*	-0.0084**	-0.0050
	(0.0031)	(0.0052)	(0.0036)	(0.0034)	(0.0048)
Regional share of R&D staff	-0.0054	-0.0091	-0.0742	0.0036	-0.0335
	(0.0466)	(0.0557)	(0.0452)	(0.0526)	(0.1008)
Observations	15395	7370	13355	12751	6940
Adj. R^2	0.020	0.039	0.028	0.020	0.032

Note: Standard errors clustered at regional level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models include time variant establishment characteristics, see Table A.7 and establishment fixed effects (equation (4)). Excluding establishments that always or never reported the respective type of innovation. *Source*: own calculations, IAB Establishment Panel, Establishment History Panel & Federal Employment Agency.

Table A.11: Impact of urbanization and localization on innovation rates conditional on time variant establishment characteristics and region fixed effects, by sector

	(1) Any innovation	(2) Radical innovation	(3) Imi- tation	(4) Improve- ment	(5) Process innovation
			Services only		
ln(Employees per km^2)	0.193	0.081	0.067	0.156	0.390
	(0.188)	(0.063)	(0.131)	(0.190)	(0.458)
Regional share of own industry	0.0042**	-0.0009	0.0003	0.0047**	0.0026
	(0.0021)	(0.0011)	(0.0022)	(0.0022)	(0.0026)
Observations	21317	21332	20746	19871	13972
Adj. R^2	0.119	0.049	0.078	0.139	0.115
			Manufacturing only		
ln(Employees per km^2)	0.137	0.031	0.260**	0.158	-0.442
	(0.140)	(0.096)	(0.116)	(0.148)	(0.478)
Regional share of own industry	0.0077***	0.0003	0.0039	0.0085***	0.0084**
	(0.0025)	(0.0021)	(0.0028)	(0.0028)	(0.0035)
Observations	14041	14045	13080	13344	9177
Adj. R^2	0.190	0.081	0.089	0.215	0.201

Note: Standard errors clustered at regional level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All models include time variant establishment characteristics reported in Table A.7 as well as region and industry fixed effects (equation (3)). *Source*: own calculations, IAB Establishment Panel, Establishment History Panel & Federal Employment Agency.

References

Alcácer, J. (2006). Location choices across the value chain: How activity and capability influence collocation. *Management Science*, *52*(10), 1457–1471.

Andersson, M., & Lööf, H. (2012). Small business innovation: firm level evidence from Sweden. *The Journal of Technology Transfer*, *37*(5), 732–754.

Andersson, R., Quigley, J. M., & Wilhelmsson, M. (2005). Agglomeration and the spatial distribution of creativity. *Papers in Regional Science*, *84*(3), 445–464.

Audretsch, D. B., & Feldman, M. P. (1996). R&D Spillovers and the Geography of Innovation and Production. *American Economic Review*, *86*(3), 630–40.

Audretsch, D. B., & Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In J. V. Henderson, & J. F. Thisse (Eds.) *Handbook of Regional and Urban Economics*, vol. 4 of *Handbook of Regional and Urban Economics*, chap. 61, (pp. 2713–2739). Elsevier.

Baptista, R., & Swann, P. (1998). Do firms in clusters innovate more? *Research Policy*, *27*(5), 525–540.

Beaudry, C., & Breschi, S. (2003). Are firms in clusters really more innovative? *Economics of Innovation and New Technology*, *12*(4), 325–342.

Bellmann, L., Crimmann, A., Evers, K., & Hujer, R. (2013). Regional Determinants of Establishments' Innovation Activities: A Multi-Level Approach. IZA Discussion Paper No. 7572, IZA - Institute of Labor Economics.

Ben Abdesslem, A., & Chiappini, R. (2019). Cluster policy and firm performance: a case study of the French optic/photonic industry. *Regional Studies*, *53*(5), 692–705.

Berliant, M., Reed III, R. R., & Wang, P. (2006). Knowledge exchange, matching, and agglomeration. *Journal of Urban Economics*, *60*(1), 69–95.

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

Bettencourt, L. M. A., Lobo, J., & Strumsky, D. (2007b). Invention in the city: Increasing returns to patenting as a scaling function of metropolitan size. *Research Policy*, *36*(1), 107–120.

Beugelsdijk, S. (2007). The regional environment and a firm's innovative performance: A plea for a multilevel interactionist approach. *Economic Geography*, *83*(2), 181–199.

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

Carlino, G., & Kerr, W. R. (2015). Agglomeration and Innovation. In G. Duranton, J. V. Henderson, & W. C. Strange (Eds.) *Handbook of Regional and Urban Economics*, vol. 5 of *Handbook of Regional and Urban Economics*, chap. 6, (pp. 349 – 404). Elsevier.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban density and the rate of invention. *Journal of Urban Economics*, *61*(3), 389–419.

Combes, P.-P., & Duranton, G. (2006). Labour pooling, labour poaching, and spatial clustering. *Regional Science and Urban Economics*, *36*(1), 1–28.

Combes, P.-P., & Gobillon, L. (2015). The Empirics of Agglomeration Economies. In G. Duranton, J. V. Henderson, & W. C. Strange (Eds.) *Handbook of Regional and Urban Economics*, vol. 5, chap. 5, (pp. 247–348). Elsevier.

De Beule, F., & Van Beveren, I. (2012). Does firm agglomeration drive product innovation and renewal? An application for Belgium. *Tijdschrift voor Economische en Sociale Geografie*, *103*(4), 457–472.

Duranton, G., & Puga, D. (2001). Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. *American Economic Review*, *91*(5), 1454–1477.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson, & J. F. Thisse (Eds.) *Handbook of Regional and Urban Economics*, vol. 4 of *Handbook of Regional and Urban Economics*, chap. 48, (pp. 2063–2117). Elsevier.

Ebert, T., Brenner, T., & Brixy, U. (2018). New firm survival: the interdependence between regional externalities and innovativeness. *Small Business Economics*, *49*, 21.

Ellguth, P., Kohaut, S., & Möller, I. (2014). The IAB Establishment Panel: Methodological essentials and data quality. *Journal for Labour Market Research*, *47*(1-2), 27–41.

Feldman, M. P., & Kogler, D. F. (2010). Stylized Facts in the Geography of Innovation. In B. H. Hall, & N. Rosenberg (Eds.) *Handbook of the Economics of Innovation*, vol. 1, chap. 8, (pp. 381–410). North-Holland.

Fornahl, D., Broekel, T., & Boschma, R. (2011). What drives patent performance of German biotech firms? The impact of R&D subsidies, knowledge networks and their location. *Papers in Regional Science*, *90*(2), 395–418.

Gerlach, H., RÃ¸nde, T., & Stahl, K. (2009). Labor pooling in R&D intensive industries. *Journal of Urban Economics*, *65*(1), 99–111.

Howell, A., He, C., Yang, R., & Fan, C. C. (2018). Agglomeration, (un)-related variety and new firm survival in China: Do local subsidies matter? *Papers in Regional Science*, *97*(3), 485–500.

Jacobs, J. (1969). *The economy of cities*. New York: Random House.

Knoben, J. (2009). Localized inter-organizational linkages, agglomeration effects, and the innovative performance of firms. *The Annals of Regional Science*, *43*(3), 757–779.

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

Malmberg, A., & Maskell, P. (2002). The elusive concept of localization economies: Towards a knowledge-based theory of spatial clustering. *Environment and Planning A*, *34*(3), 429–449.

Marshall, A. (1920). *Principles of economics*. MacMillan, London.

Moeller, K. (2018). Culturally clustered or in the cloud? How amenities drive firm location decision in Berlin. *Journal of Regional Science*, *58*(4), 728–758.

Naz, A., Niebuhr, A., & Peters, J. C. (2015). What's behind the disparities in firm innovation rates across regions? Evidence on composition and context effects. *The Annals of Regional Science*, *55*(1), 131–156.

Niebuhr, A. (2017). Die Entwicklung regionaler Disparitäten in Deutschland - Persistente Unterschiede zwischen Ost- und Westdeutschland (engl. The development of regional disparities in Germany - persistant inequalities between East and West Germany). In H. H. Eberstein, H. Karl, & G. Untiedt (Eds.) *Handbuch der regionalen Wirtschaftsförderung*, chap. A III, (pp. 1–38). Schmidt.

OECD (2005). Oslo manual: Guidelines for collecting and interpreting innovation data.

Pakes, A., & Griliches, Z. (1980). Patents and R&D at the firm level: A first report. *Economics Letters*, *5*(4), 377–381.

Pe'er, A., & Keil, T. (2013). Are all startups affected similarly by clusters? Agglomeration, competition, firm heterogeneity, and survival. *Journal of Business Venturing*, *28*(3), 354–372.

Pe'er, A., Vertinsky, I., & Keil, T. (2016). Growth and survival: The moderating effects of local agglomeration and local market structure. *Strategic Management Journal*, *37*(3), 541–564.

Peters, B. (2009). Persistence of innovation: stylised facts and panel data evidence. *The Journal of Technology Transfer*, *34*(2), 226–243.

Qian, H., & Haynes, K. E. (2014). Beyond innovation: the small business innovation research program as entrepreneurship policy. *The Journal of Technology Transfer*, *39*(4), 524–543.

Renski, H. (2011). External economies of localization, urbanization and industrial diversity and new firm survival. *Papers in Regional Science*, *90*(3), 473–502.

Sedgley, N., & Elmslie, B. (2004). The Geographic Concentration of Knowledge: Scale, Agglomeration, and Congestion in Innovation Across U.S. States. *International Regional Science Review*, *27*(2), 111–137.

Smit, M. J., Abreu, M. A., & Groot, H. L. (2013). Micro-evidence on the determinants of innovation in the Netherlands: The relative importance of absorptive capacity and agglomeration externalities. *Papers in Regional Science*, *94*(2), 249–272.

Smith, K. (2006). Measuring innovation. In J. Fagerberg, & D. C. Mowery (Eds.) *The Oxford Handbook of Innovation*, (pp. 148–177). Oxford University Press.

Srholec, M. (2010). A multilevel approach to geography of innovation. *Regional Studies*, *44*(9), 1207–1220.

Sternberg, R., & Arndt, O. (2001). The firm or the region: what determines the innovation behavior of European firms? *Economic Geography*, *77*(4), 364–382.

Tavassoli, S., & Karlsson, C. (2018). The role of regional context on innovation persistency of firms. *Papers in Regional Science*, *97*(4), 931–955.

Tödtling, F., Lehner, P., & Kaufmann, A. (2009). Do different types of innovation rely on specific kinds of knowledge interactions? *Technovation*, *29*(1), 59–71.

von Hippel, E. (2007). The sources of innovation. In *Das Summa Summarum des Management: Die 25 wichtigsten Werke für Strategie, Führung und Veränderung*, (pp. 111–120). Wiesbaden: Gabler.