



Do high local customer and supply densities foster firm growth?

Andreas Behr¹ · Christoph Schiwy² · Lucy Hong³

Accepted: 23 April 2023 / Published online: 30 May 2023
© The Author(s) 2023

Abstract We analyze the effect of local supplier and customer densities on employment growth, using a sample of 19 275 German firms. We suggest a new approach to estimating firm-specific customer and supply densities through combining input-output data from the German Federal Statistical Office and firm-level data from the Orbis database. Existing empirical analyses are primarily conducted on a regional level, so that findings may depend on the regarded regional level and only provide limited insights into the existence and effects of agglomeration on growth at the firm-level. We use kernel density estimation to avoid the arbitrariness of spatial boundaries and scales, and find that the regional agglomeration patterns vary considerably between different sectors. Our econometric analysis reveals that input supplier and customer densities have a statistically significant effect on firm growth and that the firms' age plays a crucial role in analyzing agglomeration economies.

Keywords Firm Performance · Agglomeration · Kernel density estimation · Regional Economics

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

✉ Andreas Behr
andreas.behr@vwl.uni-due.de

Christoph Schiwy
christoph.schiwy@uni-due.de

Lucy Hong
luhong@uw.edu

¹ Chair of Statistics, University of Duisburg-Essen, 45117 Essen, Germany

² University of Duisburg-Essen, 45117 Essen, Germany

³ University of Washington, Seattle, WA 98195, USA

JEL classification L25 · R11 · C10

Fördern hohe lokale Kunden- und Lieferantendichten das Unternehmenswachstum?

Zusammenfassung Wir analysieren den Effekt der lokalen Lieferanten- und Kundendichte auf das Beschäftigungswachstum auf Basis eines Datensatzes von 19.275 deutschen Unternehmen. Der vorgeschlagene Ansatz zur Schätzung der firmenspezifischen Kunden- und Lieferantendichte beruht auf der Kombination von Input-Output-Daten des Statistischen Bundesamtes und Firmendaten der Orbis-Datenbank. Bestehende empirische Analysen werden üblicherweise auf Basis verwaltungstechnisch definierten Regionen durchgeführt, deren genaue Festlegungen die Ergebnisse beeinflussen und nur begrenzte Erkenntnisse über die Existenz und die Auswirkungen der Agglomeration auf das Wachstum auf Unternehmensebene liefern. Wir verwenden eine Kerndichteschätzung zur Vermeidung der Ergebnisabhängigkeit von Verwaltungsgrenzen. Die empirischen Ergebnisse zeigen, dass die regionalen Agglomerationsmuster erheblich zwischen den verschiedenen Sektoren variieren und die Zulieferer- und Kundendichte einen statistisch signifikanten Einfluss auf das Unternehmenswachstum in Abhängigkeit vom Unternehmensalter haben.

1 Introduction

Over the last decade, a large body of literature has focused on determinants of firm growth. Convincing evidence has been provided that firm growth is affected by firm size and age, but the empirical results differ on the question of which additional firm characteristics are important determinants of firm growth (Coad 2007b).

There is a considerable literature that analyzes the relationship between financial performance and firm growth, as well as the impacts of agglomeration on firm growth. Specifically, Endogenous Growth Theory and New Economic Geography provide theoretical explanations. Since the work of Marshall (1890), it is assumed that the spatial concentration of economic activity of firms in the same sector (specialization) can foster the emergence of economies of scale, which affects the performance of firms positively. Firms may locate in proximity to other firms to benefit from a more extensive base of suppliers and customers (Glaeser and Kerr 2009, p. 636). Thereby firms can reduce their costs of obtaining inputs, often referred to as input sharing, and the costs of shipping their goods to customers. Detailed evidence of the importance of knowledge spillovers is, for example, presented by Mathias et al. (2021). Agglomeration appears to lead to knowledge gains only if unique conditions prevail, such as firm age. There is also evidence of a positive relationship between agglomeration and innovation outcomes, e.g., patents.

Existing empirical analyses are primarily conducted on a regional level¹ (e.g., city, county, or zip code level). Thus, findings may depend on the chosen regional level and provide only limited insights into the existence and effects of agglomeration on

¹ E.g., see Glaeser et al. (1992), Combes (2000), or van Oort (2007).

growth at the firm-level (de Bok and van Oort 2011). Since agglomeration economies affect regional economic growth through their impact on the performance of firms only indirectly, a firm-level agglomeration analysis is most suited to provide insight into the mechanism of agglomeration.

While there may well be effects of agglomeration, which can have effects without direct interactions of firms, e.g., positive effects of labor market pooling, these effects may be based on firm-level interactions to a larger extent. As we do not obtain direct information of supplier and customer interactions at the firm-level, sectoral data provide valuable information on input (supplier of firms' material and services for the production process) and output (customers of firms' products) relations. To utilize this information, we combine firm-level information on location and firm characteristics obtained from balance sheets and income and loss accounts, with sectoral dependencies obtained from input and output tables provided in the national account statistics published by the Federal Statistical Office of Germany (Destatis). Using this relational information, we calculate firm-specific supplier and demand densities based on surrounding firms, taking into account their relevance revealed in the sectoral input-output relations.

This paper aims to contribute to both the firm-growth and agglomeration literature by estimating a firm-growth model, which regards firm-specific attributes, as well as agglomeration effects based on supplier and customer densities in Germany.

Although there is literature that examines the effect of input-output linkages (as well as other Marshallian agglomeration mechanisms) on coagglomeration, which is the tendency of two industries to locate in proximity (e.g., see Ellison et al. 2010 or Aleksandrova et al. 2020), this paper differs from these others by using kernel densities to calculate supplier and customer densities specific to each individual firm's location, thereby avoiding the arbitrariness of spatial boundaries and scales.

The paper is structured as follows. In the next section, we provide a literature review. In the third section, we present our database and the operationalization of supplier and customer densities. Section four includes descriptive evidence and regression results. We discuss our findings and provide conclusions in the last section.

2 Literature Overview

In our analysis, we examine the effect of the agglomeration of supplier and customer firms at the firm's specific location, on its employment growth. Hence, we discuss the firm growth literature as well as the agglomeration literature that is relevant to this specific linkage.

Firm age and firm size are the most frequently investigated determinants of firm growth in empirical studies. Age and size may be correlated, since large firms are mostly older than smaller firms (Coad 2007b, p. 51). Reviewing previous research on the growth of firms, Coad (2007b, p. 18) reveals that age affects firm growth negatively. Similar findings are also found by Harhoff et al. (1998) and Audretsch and Dohse (2007), who analyze employment growth using a dataset of German firms. Concerning firm size (total assets), Harhoff et al. (1998), Audretsch and

Dohse (2007), and Hoogstra and van Dijk (2004) report a negative impact of firm size on employment growth.

Another strand of the literature focusses on the link between financial performance and firm growth. For instance, previous sales growth may spur growth, since investors may consider firms with high growth in sales as secure investments (Levratto et al. 2010, p. 8). In this context, Fuertes-Callen and Cuellar-Fernandez (2019) analyze the impact of profitability, measured as the return on assets (earnings before interest and taxes divided by total assets), on Spanish employment, as well as on sales growth using static and dynamic panel models. They discover a positive effect of past profits on employment growth, but a negative effect on sales growth, suggesting that Spanish firms do not tend to invest in growth. This result is in line with Coad (2007b), who states that although financial performance is often found to be statistically significant, its impact on growth is small.

The financial structure of a firm may also influence its growth. A high debt ratio may indicate lower financial constraints, but could also increase the risk of future financial constraints and solvency, whereas a high fixed assets ratio may lead to low flexibility and the need to cover recurring asset expenses. However, in empirical research, it remains unclear whether the financial structure is an essential determinant of firm growth. For example, Becchetti and Trovato (2002, p. 294) find that high-leverage firms tend to experience stronger growth than low-leverage firms, whereas, in the study of Fuertes-Callen and Cuellar-Fernandez (2019, p. 96), the debt ratio does not have a statistically significant influence on employment growth.

A wide range of studies have analyzed the relationship between regional growth and agglomeration economies using aggregated regional data, such as zip-code-sectors or municipal-sectors (e.g., Glaeser et al. 1992; Combes 2000; van Oort 2007). Agglomeration economies may arise from various types of economic structure, for instance, the spatial concentration of economic activity in the same sector (specialization), in diverse sectors (diversity), or competition. Although there is no agreement on which type of economic structure leads to the emergence of agglomeration economies, many empirical studies have indicated a link between agglomeration economies and the uneven regional distribution of economic activity and growth (de Bok and van Oort 2011, pp. 5–6).

Mathias et al. (2021) review 42 studies on the relationship between agglomeration, which was operationalized by aggregated regional indices, and firm performance. They find that agglomeration is associated with innovation benefits, but the level of agglomeration depends on unique conditions such as firm age or the institutional environment. Mathias et al. (2021, p. 436) conjecture that the positive relationship between the degree of agglomeration and financial performance outcomes will be stronger in younger firms than in older ones. In contradiction to their supposition, they find that older firms achieve greater financial performance (Mathias et al. 2021, p. 441).

However, the spatial scale composition may affect the regional analysis (see the Modifiable Areal Unit Problem (MAUP)²). Furthermore, these studies only provide

² For instance, Burger et al. (2010) provide some insight into the impact of the spatial scale composition on the relationship between agglomeration externalities and regional growth.

limited information on the relationship between agglomeration and firm performance.

These limitations may be due to two reasons: On the one hand, agglomeration economies only indirectly affect regional economic growth through their impact on the performance of firms (de Bok and van Oort 2011, pp. 5–6). On the other hand, these studies often use economic structure indicators, which are hypothesized as fostering the emergence of agglomeration economies, to examine agglomeration forces. For example, the location quotient, a sectoral employment share indicator, is used to measure specialization. However, in theory, not the economic structure per se, but agglomeration mechanisms lead to externalities, potentially fostering firm performance. Moreover, specific economic structures such as specialization, may also generate negative effects, e.g., congestion effects (Henderson et al. 1995) or vulnerability to sector-specific shocks (Duranton and Puga 2004), which is why measuring economic structures may only be second-best.

Over the last two decades, several scholars have focused on sources of agglomeration.³ Using U.S. manufacturing firm data, Rosenthal and Strange (2001) find that all three Marshallian agglomeration mechanisms (labor market pooling, input sharing, knowledge spillovers) positively affect the emergence of agglomeration economies.

Ellison et al. (2010) examine the relative importance of the Marshallian agglomeration mechanisms using OLS regressions based on firm-level manufacturing and input-output accounting data. Regressing two different coagglomeration indices, which measure the tendency of two industries to locate in proximity to one another, on variables measuring the extent of input-output linkages, separate input and output effects, labor market pooling, and technological flows, they find a positive statistically significant effect for all three Marshallian mechanisms on the tendency to coagglomerate. However, of these three mechanisms, input-output linkages are particularly important.

Jofre-Monseny et al. (2011) and Aleksandrova et al. (2020) also provide supportive evidence of the relevance of input-output linkages by analyzing the importance of the three Marshallian agglomeration mechanism concerning the creation of firms across cities (between-cities) and across municipalities within large cities (within-cities) in Spain and Russian manufacturing firms, respectively. Since the use of a coagglomeration index as the dependent variable may lead to simultaneity, i.e., customer-supplier relationships may be the result of coagglomeration, Jofre-Monseny et al. (2011) employ a Poisson regression to model the count of new firms. To define customer-supplier relations, data from the Catalan Input-Output Table is used to create variables measuring local employment in industries that are the main input supplier or customers, respectively, of an industry. Their results indicate that only input sharing and labor market pooling constitute important agglomeration mechanisms in the between-cities analysis, but all three mechanisms are relevant in the within-cities analysis.

³ A review of earlier agglomeration literature can be found in Rosenthal and Strange (2004), McCann and Folta (2009), or Combes and Gobillon (2015). See Nakamura and Paul (2019) for a review on agglomeration measurements and sources of agglomeration.

Also addressing the micro-foundations of observed location patterns at state, county, and zip code level, Kolko (2010) focuses on the question of why services industries are more urbanized, but less highly concentrated than manufacturing industries. The author finds that services industries rely less on natural resources, and occupational specialization does not have a statistically significant effect on coagglomeration. Services industries also tend to urbanize, in order to minimize transport costs for their outputs.

To study sectoral localization, a strand of literature uses a nonparametric approach developed by Duranton and Overman (2005), which compares the distribution of pairwise distances between establishments in an industry to that of randomly chosen establishments. Regarding the colocalization patterns between vertically-linked industries, Duranton and Overman (2008) discover that on small spatial scales, establishments tend to locate closer to the same industry establishments, while on a greater spatial scale (around 150 km), they tend to locate closer to vertically-linked industries. Klier and McMillen (2008) find that U.S. auto supplier plants tend to locate in areas with good highway access, close to Detroit, and in proximity to assembly plants. Recent evidence from Goldman et al. (2019) suggests that R&D, administrative, and production workers are the most concentrated occupation groups for manufacturing industries. This result is in line with the argument that knowledge spillovers and labor market pooling lead to a high concentration of R&D and administrative workers.

The literature also provides empirical insights into the sources of agglomeration about production and productivity. For instance, Feser (2002), Rigby and Essletzbichler (2002), and Greenstone et al. (2010) estimate production functions using plant-level data and found evidence for an influence of the traditional Marshallian sources of agglomeration in the U.S.

In conclusion, while many studies have examined the impact of agglomeration on regional growth, the empirical literature focusing on individual firms is rather limited. Existing studies have regressed (co-) agglomeration indexes or (total factor) productivity on Marshall's three sources of agglomeration. In particular, to gain deeper insight into the relationship between agglomeration and regional growth, it is necessary to analyze the effects of the micro-foundations of agglomeration economies on regional growth, in addition to the economic structures conducive to agglomeration. Therefore, focusing on regional agglomeration effects, considering the surrounding supplier and customer locations information on the individual firm's growth seems to be a promising research strategy.

3 Data and variable construction

3.1 Data

For the analysis, we use firm-level information from the Orbis database generated by "Bureau van Dijk" (BvD). This database includes companies of all sizes, but has a focus on private companies. Orbis contains balance sheet and income statement firm-level data and in addition, Orbis also provides data from the notes to the finan-

cial statements, such as the number of employees, or metadata such as information about a company's industry. The database also includes the firm's address, which was transformed into geographic coordinates (longitude and latitude).

The data quality requirement for firms to provide information on the supplier and customer densities is relatively low, as we only need information on the firm location and their sales in 2013. Firms included in the final regression analysis have to fulfill more restrictive criteria, as we need valid information on all the variables included in the regression model. Hence, we constructed two datasets, one with fewer variables and less stringent selection criteria for the estimations of the input supplier and customer densities, and one more restricted dataset for the core analysis of growth determinants.⁴ The number of observations in the data set for estimating supplier and customer densities is 21 053, whereas the number of observations in the data set finally used for the regression analysis is 19 275.

We use input-output-accounts provided by the German Federal Statistical Office (Statistisches Bundesamt) to measure the input supplier and customer linkages between sectors. Sector information in the Orbis dataset is provided as a NACE code (Eurostat 2008), while the aggregation to 72 sectors in the specific advanced input matrix is based on CPA (Classification of Products by Activity). According to the Statistisches Bundesamt (2008, pp. 48–49), the structure of CPA generally corresponds to that of NACE up to the level "class".

Starting with the advanced input matrix (72×72) for domestic production (Table 2.3 provided in Statistisches Bundesamt 2017, pp. 60–71), we aggregated the 72 sectors to 12, using the conversion table provided in Behr and Rohwer (2019, p. 270) (see also Table 4). For the Orbis dataset, we first converted NACE to the 72 sectors, as given in the advanced input matrix, and then aggregated further to 12 sectors.

We compare our data (dataset used for the regression analysis) with census data from the regional database (Regionaldatenbank Deutschland 2020) provided by the German Federal State Statistical Offices to assess the quality of the dataset. German census data is classified according to WZ 2008⁵, which structure corresponds to that of NACE according to Statistisches Bundesamt (2008, pp. 47–48). According to Table 5 and Table 6, our data reflects the employment share per sector and federal state of the census data. In Table 5, we find similar employment shares per sector for both datasets, and from Table 6 we find that employment shares per federal state are similarly distributed.

3.2 Variable construction

In this subsection, we describe the definition of variables obtained directly from firm-level balance sheets and income and loss statements.

⁴ The selection criteria to ensure certain data quality for both data sets are listed in Table 3 in the appendix.

⁵ Classification of economic activities, 2008 edition (WZ 2008).

Employment growth. Using job creation as the general objective from a socio-economic perspective, in line with Illy et al. (2009), we use the employment growth indicator

$$lgr_i = \ln(emp_{i,2017}) - \ln(emp_{i,2013}) \quad (1)$$

as a proxy for economic growth, where emp_i represents (total) employment (number of employees) in firm i .

Firm age. Several studies have shown that young firms tend to grow faster than older ones, e.g., Harhoff et al. (1998). We define the age of the firm according to the available information in the Orbis data set in decades:

$$age_i = (\text{age in days})_{i,2013} / 3650. \quad (2)$$

The more dynamic growth process of younger companies tends to have a relatively high variance, see, e.g., Coad (2007b), Dunne et al. (1989) and Boeri and Cramer (1992).

Firm size. We define a firm's size by the natural logarithm of its total assets

$$ltoas_i = \ln(\text{total assets})_{i,2013}. \quad (3)$$

A firm's age and size have been found to be important in firm-growth-related studies. Since larger firms are usually older than smaller ones, there is possibly a correlation between these two predictors (Coad 2007b, p. 18 and p. 51). Furthermore, small firms are under constant pressure to grow in order to reduce costs to the same level as other (larger) competitors (Coad 2007b, p. 51).

A specific minimal size is essential in fields with high fixed costs, whereas a relatively small size increases the firm's flexibility (Schreyer 2000, p. 13). Consequently, large firms have longer planning horizons and more long-term investments, which pay off over several years and lead to a higher autocorrelation of growth (Coad 2007a, p. 74). In addition national legislation is often firm-size-specific. For instance, larger firms in certain countries have higher firing costs and have to pay higher taxes. On the other hand, they also have greater lobbying power, which might facilitate growth (Coad 2007b, p. 27 and p. 78–79 and Schreyer 2000, p. 7).

There seems to be no consensus in the literature as to whether a small or a large size facilitates high growth. For instance, Harhoff et al. (1998) find that a larger size leads to lower growth rates, whereas Henrekson and Johansson (2010, p. 1) find the opposite.

Sales growth. We consider the change in sales

$$lsa_i = \ln(sales_{i,2013}) - \ln(sales_{i,2012}) \quad (4)$$

as it can be considered a main determinant for firm growth to meet the rising demand for its product by enlarging its employment.

The debt ratio. The debt ratio is defined as follows:

$$dr_i = \frac{(\text{total liabilities})_{i,2013}}{(\text{total assets})_{i,2013}} \quad (5)$$

It measures a firm's leverage, indicating the borrowed share of a firm's funding. The remaining financing originates from past retained profits or has been introduced by the shareholders (Albrecht et al. 2007, p. 476; Penner 2004, p. 218). One important motivation to include dr has been put forward by Lopez-Garcia and Puente (2012, p. 1036–1039). They argue that firm growth requires financing and that a high dr might lead to future financing constraints and an inability to realize all reputedly profitable projects. In their study, dr was found to have a significant non-linear influence when not controlling for firm-specific time-invariant factors. The importance of dr for firm growth was also confirmed by Fagiolo and Luzzi (2006, p. 33), Becchetti and Trovato (2002, p. 294) and Levratto et al. (2010, p. 10). Lopez-Garcia and Puente (2012, p. 1031) found no effect of dr . One possible reason for doubt is that, especially in cases of start-ups, not bank credits but risk capital and internal finance are important funding sources (Lopez-Garcia and Puente 2012, p. 1038).

The fixed assets ratio. The fixed assets ratio

$$far_i = \frac{(\text{fixed assets})_{i,2013}}{(\text{total assets})_{i,2013}} \quad (6)$$

indicates the degree of capital commitment. High values imply low flexibility and constant pressure to keep capacity utilization high, so as to cover the recurring assets' expenses (e.g., for energy and maintenance). We assume that a high far might motivate a firm to grow in order to spread its high fixed costs over a greater number of products (Schneider and Lindner 2010, p. 317). Levratto et al. (2010, p. 5) added that excessively high fixed costs could be a severe threat to future growth.

3.3 Firm-level supplier and customer densities

In this section, we describe our approach of combining firm-level and input-output data to obtain firm-level estimates of supplier and customer densities.

3.3.1 Sector densities

In order to calculate input supplier and customer densities, we first calculate n_s weighted sector densities d_i^s for each firm i at each firms' coordinates. d_i^s is the density estimate, using all firms from sector s with $s = 1, \dots, n_s$ (weighted with their amount of sales in 2013), evaluated at the firm's location $\{\text{longitude}_i, \text{latitude}_i\}$.⁶

⁶ We use n to indicate the number of firms and n_s for the number of sectors.

Following the notation of Wand and Jones (1995, p. 90), the multivariate density estimator for a d -dimensional random sample $\mathbf{X}_1, \dots, \mathbf{X}_n; \mathbf{X}_i \in \mathbb{R}^d$ with density f can generally be written as⁷

$$\widehat{f}(\mathbf{x}; \mathbf{H}) = n^{-1} \sum_{i=1}^n K_{\mathbf{H}}(\mathbf{x} - \mathbf{X}_i), \tag{7}$$

with

$$K_{\mathbf{H}} = |\mathbf{H}|^{-1/2} K\left(|\mathbf{H}|^{-1/2} \mathbf{x}\right), \tag{8}$$

where $\mathbf{x} \in \mathbb{R}^d$, \mathbf{H} is a $d \times d$ bandwidth matrix (symmetric positive definite) and $K(\cdot)$ is a d -variate kernel function with $\int K(\mathbf{x})d\mathbf{x} = 1$. In this paper, the standard normal density,

$$K(\mathbf{x}) = (2\pi)^{-d/2} e^{-(1/2)\mathbf{x}'\mathbf{x}}, \tag{9}$$

was used as kernel function. To account for weights, we can modify Eq. (7) towards

$$\widehat{f}(\mathbf{x}; \mathbf{H}; \mathbf{y}) = n^{-1} \sum_{i=1}^n \omega_i K_{\mathbf{H}}(\mathbf{x} - \mathbf{X}_i), \quad \text{with} \quad \omega_i = \frac{y_i n}{\sum_{i=1}^n y_i}, \tag{10}$$

where \mathbf{y} is a vector with weights (sales).

We use the same bandwidth matrix \mathbf{H} for all twelve sector-specific density estimations to obtain an equal surrounding area. The bandwidth is estimated by a plug-in bandwidth selector, using all firm locations, regardless of the specific sector. Calculating the weighted densities with the fixed bandwidth matrix, we obtain a vector $\mathbf{d}_i = (d_i^1, \dots, d_i^s, \dots, d_i^{n_s})'$ for each firm i containing the sectoral densities.

3.3.2 Taking into account sectoral supplier and customer interrelations

To construct input supplier and customer densities, we combine input-output data from the German Federal Statistical Office and firm-level data from the Orbis database. Note that we have only locational information on intermediate goods, and therefore, the obtained customer density has to be calculated based on surrounding firms buying intermediate products for their own production process.

Given a symmetric $n_s \times n_s$ advance input matrix for n_s sectors

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & \dots & a_{1,n_s} \\ \vdots & \ddots & \vdots \\ a_{n_s,1} & \dots & a_{n_s,n_s} \end{pmatrix}, \tag{11}$$

where j ($j = 1, \dots, n_s$) is an index for rows and k ($k = 1, \dots, n_s$) is an index for columns, $a_{j,k}$ is the value of goods of sector j , which are used as the input for

⁷ Here we present a general notation. Note that in this paper, \mathbf{X} has two dimensions, longitude and latitude.

sector k . Note that all entries of the matrix are measured in Euros, thus allowing for aggregation.

Thus, the j -th row sum $\sum_{k=1}^{n_s} a_{j,k}$ represents the value of all advanced inputs that sector j delivers to its “customers”, analogously the column sum $\sum_{j=1}^{n_s} a_{j,k}$ is the value of all advanced inputs that sector k uses for production.

To account for the different sector shares in the economy, we define two sector-specific $n_s \times 1$ weight vectors

$$S^{\text{sup}} = \frac{1}{\sum_{j=1}^{n_s} \sum_{k=1}^{n_s} a_{j,k}} \begin{pmatrix} \sum_{j=1}^{n_s} a_{j,1} \\ \vdots \\ \sum_{j=1}^{n_s} a_{j,n_s} \end{pmatrix} \tag{12}$$

and

$$S^{\text{cus}} = \frac{1}{\sum_{j=1}^{n_s} \sum_{k=1}^{n_s} a_{j,k}} \begin{pmatrix} \sum_{k=1}^{n_s} a_{k,1} \\ \vdots \\ \sum_{k=1}^{n_s} a_{k,n_s} \end{pmatrix} \tag{13}$$

where S^{sup} is the weight vector for the input supplier density and S^{cus} the weight vector for the customer density.

Assume that firm i at whose location the density is to be estimated belongs to sector s . We use the s -th column of the input-output matrix to define the weight vector for its sectoral suppliers as

$$W_s^{\text{sup}} = \frac{1}{\sum_{j=1}^{n_s} a_{j,s}} \begin{pmatrix} a_{1,s} \\ \vdots \\ a_{n_s,s} \end{pmatrix} \tag{14}$$

where W_s^{sup} is the weight vector for the input supplier density for firm i located in sector s .

Accordingly, we use the s -th row of the input-ouput matrix to define the weight vector for its sectoral customers as

$$W_s^{\text{cus}} = \frac{1}{\sum_{k=1}^{n_s} a_{s,k}} \begin{pmatrix} a_{s,1} \\ \vdots \\ a_{s,n_s} \end{pmatrix} \tag{15}$$

where W_s^{cus} is the weight vector for customer density. Note that all firms located in the same sector are therefore assumed to have the identical sectoral supplier and customer weighting scheme for their surrounding firms. Note also the two different functions of the weighting schemes S^{sup} , S^{cus} and W_s^{sup} , W_s^{cus} . As all sectoral densities have an identical volume of 1 by definition, S^{sup} and S^{cus} are used to account for the different shares of the sectors in total supply and demand. W_s^{sup} and W_s^{cus} account for the firm-specific supply and customer relations based on sectoral information.

Using the weighted density estimate, we can calculate the input supplier and customer density for a firm i belonging to sector s as

$$d_i^{\text{sup}} = (\mathbf{S}^{\text{sup}} \circ \mathbf{W}_s^{\text{sup}})' \mathbf{d}_i \quad \text{and} \quad d_i^{\text{cus}} = (\mathbf{S}^{\text{cus}} \circ \mathbf{W}_s^{\text{cus}})' \mathbf{d}_i, \quad (16)$$

where \circ is the Hadamard product.

3.3.3 Regression equations

Since the emergence of agglomeration economies appears to be affected by unique conditions, such as firm age (Mathias et al. 2021), we conjecture that nascent firms may benefit from higher supply and customer densities. Based on the age of the firms, we form an indicator function I_a to group the firms into “old” and “young” ones (1, if a firm belongs to the group of “old” companies, 0 otherwise), using the median age of the firms as threshold.

We estimate the following regression

$$\begin{aligned} \ln r_i = & \beta_0 + \sum_{j=1}^{n_s-1} \beta_j S_{i,j} + \beta_{12} \text{age}_i + \beta_{13} \text{ltoas}_i + \beta_{14} \text{lsa}_i + \beta_{15} \text{dr}_i + \beta_{16} \text{far}_i \\ & + \beta_{17} d_i^{\text{sup}} + \beta_{18} I_a d_i^{\text{sup}} + \beta_{19} d_i^{\text{cus}} + \beta_{20} I_a d_i^{\text{cus}} + \epsilon_i. \end{aligned} \quad (17)$$

$S_{i,1}, \dots, S_{i,11}$ take the value 1 if i belongs to sector S_j and 0 otherwise, using sector twelve (here: construction) as the reference sector, ϵ_i is an independent and identically distributed (i.i.d.) error term and i is the firm index ($i = 1, \dots, n$).

4 Empirical findings

4.1 Descriptive results

For the regression dataset with 19 275 observations, we find (see Table 7) that most firms in the dataset belong to the trade and finance sectors (30.78% and 21.08%), and the least firms belong to vehicles and mining (0.6% and 0.43%). Looking at the number of employees per sector, we find that most employees work in the sectors finance and trade (41.39% and 18.53%) and in sectors agriculture and mining (0.52% and 0.27%) the least.

For the density estimation dataset (21 053 observations), we show in Figs. 1 and 2 the sales-weighted density estimations for each sector. The darker the area, the higher the estimated weighted density at this region.

Firms from the agriculture sector generally seem to be located in the new federal states, with some larger hotspots in the old federal states. The sectors of mining, chemistry, machinery, and processing are primarily concentrated in the old federal states. For the other sectors, the weighted densities appear to be more evenly distributed. While firms from the food and vehicles sectors show a rather spotty

Fig. 1 Densities for sectors 1–6. **a** agriculture, **b** mining, **c** food, **d** chemistry, **e** machinery, **f** vehicles

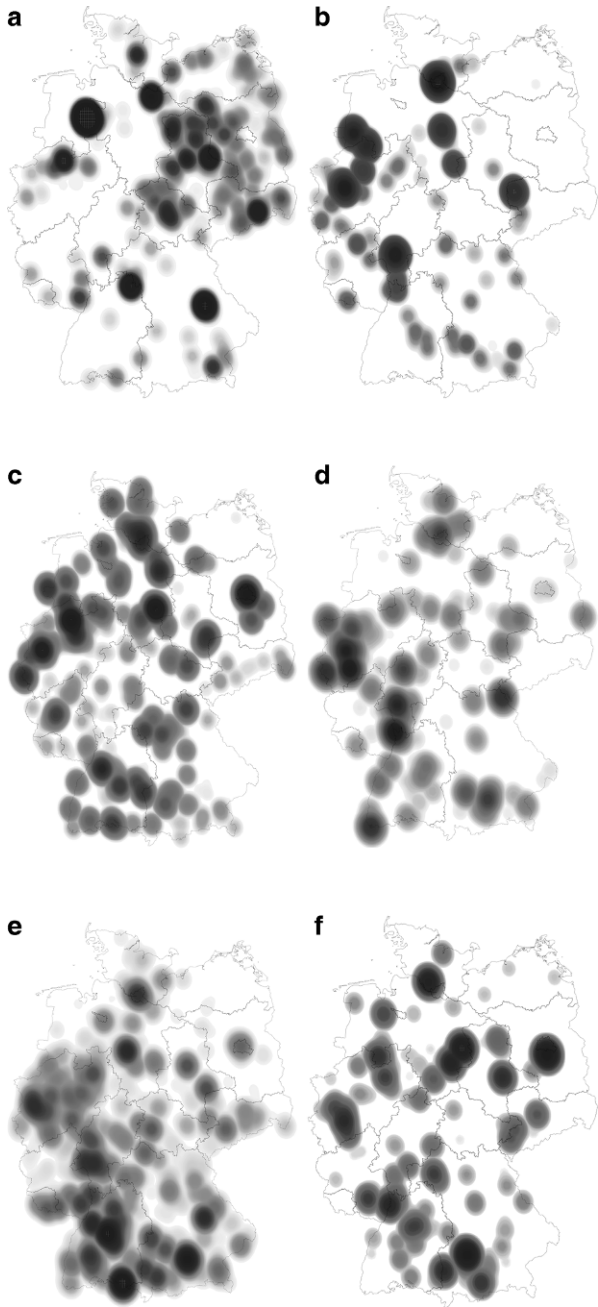


Fig. 2 Densities for sectors 7–12. **a** processing, **b** energy, **c** construction, **d** trade, **e** finance, **f** services

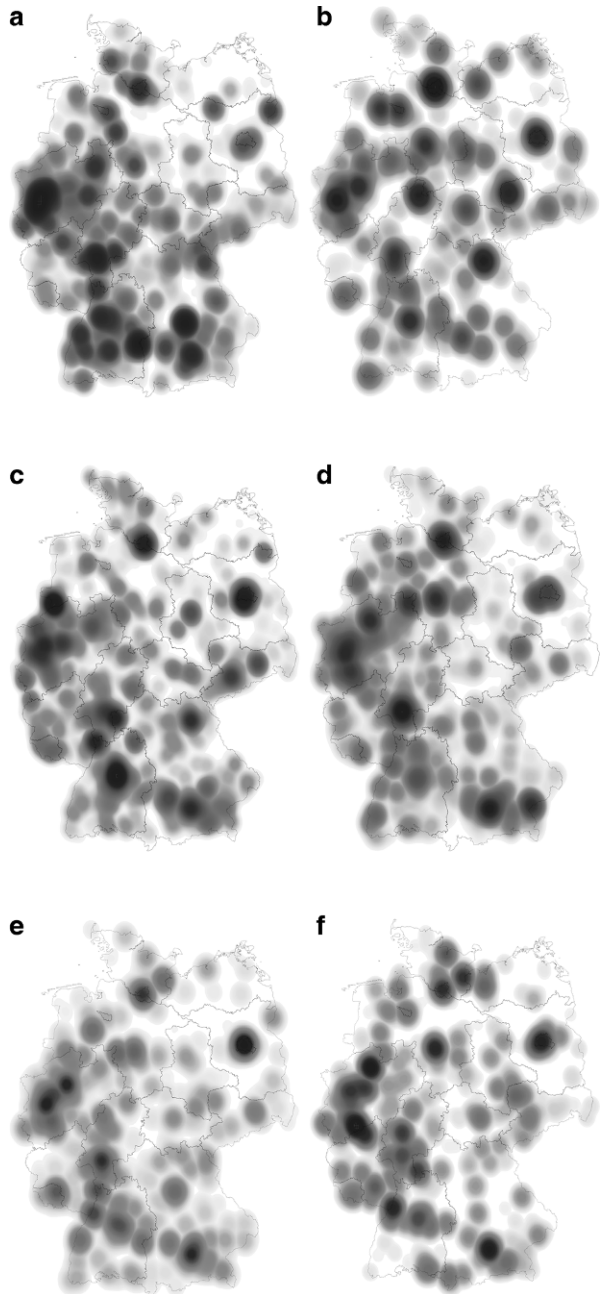


Table 1 Descriptive statistics for the variables used in the regressions

Statistic	<i>lgr</i>	<i>age</i>	<i>ltoas</i>	<i>lsa</i>	<i>dr</i>	<i>far</i>	<i>sden</i>	<i>dden</i>
Min	-2.9957	0.0359	3.1355	-0.9985	0.0064	0.0000	0.0000	0.0000
Median	0.0353	2.0934	7.2584	0.0131	0.6538	0.2174	0.0045	0.0040
Mean	0.0823	2.8061	7.5858	0.0302	0.6442	0.2934	0.0191	0.0185
Max	2.9741	52.5597	16.8416	0.9995	3.8261	1.0000	0.2208	0.2269
Sd	0.4028	2.7274	1.6982	0.1704	0.3010	0.2492	0.0303	0.0315

distribution in space, whereas energy, construction, trade, finance and services seem to be more evenly distributed.

In Table 1, descriptive statistics for the regression variables are presented. We find that the distribution of the approximate percentage growth rates for employees *lgr* is right-skewed with a mean of 8.23% and a median of 3.53%. The median firm age is about 21 years, whereas the youngest firm in the dataset has existed for less than a year, and the oldest firm for more than 525 years. Firm size ranges significantly between €23k and €20,617,000k, for which the amount of total assets of the median firm is €1,420k. For approximate growth in sales, we find that more than half of the firms have a positive growth of more than about 1.31%. Mean leverage is about 65% and the mean of the fixed asset ratio for all sectors is 29.34%, but varies considerably between sectors.

4.2 Regression results

To empirically determine the effects of supplier and customer relationships, we estimate regression Eq. (17) and show the results in Table 2.⁸

The estimated input supplier and customer densities both have a statistically significant but opposing effect on firm growth for younger firms. In theory, firms would benefit from a dense supplier and customer environment, reducing costs for obtaining inputs from suppliers and shipping products to their customers. For older firms, the impacts point in the same direction, but are much smaller, thus indicating that younger firms are much more sensitive to agglomeration economies. In general, employment growth for German firms seems to benefit from a high input supplier density, whereas a high customer density seems to hamper growth.

For size (*ltoas*) and age, we find a significant negative relationship with firm growth, which is in line with previous literature. While Fuertes-Callen and Cuellar-Fernandez (2019) found a negative effect of sales growth on employment growth for Spanish manufacturing firms, we see a significant positive overall effect for German firms.

Compared to sector construction, firm growth is more extensive for firms belonging to a different sector, except for mining and agriculture, whereas only the effect of the sector dummy for agriculture is statistically significant.

⁸ Given that conley standard errors are used, Table 8 shows a robustness check for multiple distance cutoffs.

Table 2 Regression results (using sector construction as the reference sector)

Coef./Fit		Coef.	
Constant	0.1582*** (0.0226)	processing	0.0225*** (0.0079)
Mining	-0.0395 (0.0360)	age	-0.0085*** (0.0007)
Chemistry	0.0952*** (0.0236)	ltoas	-0.0114*** (0.0024)
Services	0.0968*** (0.0148)	lsa	0.2857*** (0.0131)
Energy	0.0408** (0.0207)	dr	-0.0059 (0.0126)
Vehicles	0.0118 (0.0300)	far	0.0013 (0.0157)
Finance	0.0490*** (0.0102)	d^{sup}	1.2688*** (0.4297)
Trade	0.0364*** (0.0069)	$I_a d^{\text{sup}}$	-1.0316* (0.5706)
Agriculture	-0.1040*** (0.0224)	d^{cus}	-1.3910*** (0.3851)
Machinery	0.0242* (0.0144)	$I_a d^{\text{cus}}$	0.9494* (0.5061)
Food	0.0741*** (0.0208)		
Obs.	19275		
R ²	0.0276		
Adj. R ²	0.0266		
Wald (F)	27.3308***		

Notes: Conley standard errors estimated with a distance cutoff of 250 km are given in parentheses.

* Significance at the 10% level.

** Significance at the 5% level.

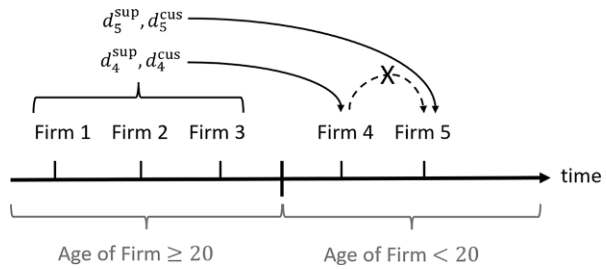
*** Significance at the 1% level.

To measure the effect of the financial structure of a firm on employment growth, we included the debt ratio (dr) and the fixed asset ratio (far). For our dataset, both variables have no statistically significant effect on firm growth.

To address concerns of potential endogeneity of agglomeration, i.e. new firms may cluster close to fast growing, overperforming firms, thereby reversing the causality and to analyze the robustness of our regression results in this respect, we perform the following analysis.

First, we calculated the supplier and demand densities based on a subsample consisting only of firms above or equal the age of 20. Second, the regression (Eq. 17) is now based on the subsample consisting only of firms younger than 20. Hence, all firms considered for estimating the densities chose their location prior to the existence of the younger firms used in the regression. The grouping towards younger

Fig. 3 The concept of the identification strategy using a simple example



and older firms is now based on the median age of about 12 years for the subsample of younger firms.

Fig. 3 presents our identification approach using a simple example of five firms. We divided our sample into two groups based on firm age. Supplier and customer densities for firms 4 and 5 ($d_4^{\text{sup}}, d_5^{\text{sup}}, d_4^{\text{cus}}, d_5^{\text{cus}}$) are then calculated using the subsample of firms above or equal to the age of 20 (firms 1 to 3). This ensures that agglomeration effects affect the firms in the subsample for the regression analysis (firms 4 and 5) in only one direction. Thus, firms 1–3 can affect firms 4 and 5, but there is no effect of firm 4 (a hypothetical firm with above-average performance) on firm 5.

We present these additional regression results in Table 9 and provide a robustness check in Table 10. The regression results show that the estimated regression coefficients resemble the coefficients in the original analysis very closely. The significances of the hypothesis tests also resemble the original results. Due to roughly halving the number of observations by considering only younger firms, of course the standard errors are slightly larger as are the p-values. Based on this regression findings we conclude that our original analysis indeed is not biased by potential endogeneity.

5 Conclusion

Since agglomeration economies affect regional growth mediated through firm growth, we provide a firm-level approach to analyze the effects of firm-specific supplier and customer densities. Estimating these densities is enabled by combining sectoral supplier and customer interrelations of input-output tables with firm-level information obtained from balance sheets and profit and loss accounts from the Orbis database.

The use of kernel density estimates allows us to avoid the use of arbitrary spatial boundaries and scales and to use individual firms as the unit of analysis. Analyzing the effects of supplier and customer relations for firms is more appropriate than the traditional analysis based on region-sector combinations.

We observe that regional agglomeration patterns differ substantially between sectors. Our econometric analysis, also considering firm-specific covariates such as firm size, age, and sales growth, shows that both input supplier and customer densities have a statistically significant effect on firm growth. As the effects differ in

size for younger and older firms, firm age seems to be an important determinant when analyzing agglomeration economies. According to theory, a dense supplier and customer environment may be expected to reduce the costs of obtaining inputs and delivering goods to customers. Nevertheless, high regional densities may also implicate disadvantages, such as traffic congestion, high competitive pressure, or high rents. Based on our German data set, we find that a dense supplier environment contributes to firm growth, whereas a dense customer environment seems rather to hamper employment growth. One factor that is difficult to capture in data is that local policies may be designed to incentivize or discourage firms from locating in areas with a high supplier or customer density. Those policies may eventually lead to a regionally suboptimal relationship between suppliers and customers.

As our analysis might potentially be prone to the problem of endogeneity, i.e., high growth may as well attract other firms resulting in higher agglomeration, we provide an additional analysis using only agglomeration information obtained from a subsample before the period of analysis. This analysis confirms the results of our main analysis, indicating the absence of endogeneity bias.

In summary, using the firm-level supplier and customer densities allows for a more profound understanding of the relationship between agglomeration mechanisms and specific firm characteristics. Future research might explore how the impacts of dense supplier and customer environments vary depending on a firm's stage of development, market power, or size. Additionally, the interaction of agglomeration and firm age might be analyzed more closely in further research.

6 Appendix

6.1 Tables

Table 3 Validity of the dataset

Dataset	Variable	Criteria
For density calculation		
	Period per fiscal year	12 months
	Accounting practice	local GAAP
	Firm status	active
	Total assets in 2013	$\geq \text{€}10\,000$
	Sales in 2012 and 2013	$\geq \text{€}10\,000$
	Number of employees in 2013 and 2017	≥ 1
Descriptives and regression		
	Period per fiscal year	12 months
	Accounting practice	local GAAP
	Firm status	active
	Total assets in 2013	$\geq \text{€}10\,000$
	Sales in 2012 and 2013	$\geq \text{€}10\,000$
	Non current liabilities in 2013	$\geq \text{€}0$
	Current liabilities in 2013	$\geq \text{€}0$
	Total liabilities in 2013	$> \text{€}0$
	Number of employees in 2013 and 2017	≥ 1
	Age in days	≥ 1
	<i>lgr</i>	$\geq -3 \ \& \ \leq 3$
	<i>lsa</i>	$\geq -1 \ \& \ \leq 1$
	<i>dr</i>	≤ 4

Table 4 Sector aggregation

Sector	Name – english	Name – german	aggregated 72
1	Agriculture	Land	1–3
2	Mining	Bergbau	4–6
3	Food	Nahrung	7
4	Chemistry	Chemie	13, 14
5	Machinery	Maschinen	22–24
6	Vehicles	Fahrzeuge	25–26
7	Processing	Verarbeitung	8–12, 15–21, 27, 28
8	Energy	Energie	29–32
9	Construction	Bau	33–35
10	Trade	Handel	36–44
11	Finance	Finanzen	45–61
12	Services	Dienste	62–72

Table 5 Comparison of employee shares per sector for Germany in 2017. We compare the Orbis data with census data from the regional database (Regionaldatenbank Deutschland 2020)

NACE	WZ08	Orbis	Regional DB
01–03	A	0.0054	0.0139
05–39	B–E	0.2590	0.1848
41–43	F	0.0312	0.0560
45–98	G–T	0.7043	0.7453

Table 6 Comparison of employee shares per federal state for Germany in 2017. We compare the Orbis data with census data from the regional database (Regionaldatenbank Deutschland 2020)

Federal State	Orbis	Regional DB
Baden-Wuerttemberg	0.1736	0.1413
Bavaria	0.1495	0.1701
Berlin	0.0416	0.0443
Brandenburg	0.0139	0.0251
Bremen	0.0064	0.0096
Hamburg	0.0332	0.0284
Hesse	0.1492	0.0779
Mecklenburg-Vorpommern	0.0069	0.0169
Lower Saxony	0.0595	0.0917
North Rhine-Westphalia	0.2438	0.2130
Rhineland-Palatinate	0.0251	0.0456
Saarland	0.0244	0.0120
Saxony	0.0238	0.0462
Saxony-Anhalt	0.0131	0.0227
Schleswig-Holstein	0.0246	0.0316
Thuringia	0.0114	0.0236

Table 7 Number of firms and employees per sector for the regression dataset

Sector	Number of firms	Number of employees
Construction	2931	74899
Mining	83	6425
Chemistry	211	61866
Services	834	248050
Energy	559	79799
Vehicles	116	27270
Finance	4063	1002063
Trade	5932	448602
Agriculture	275	12482
Machinery	1334	208987
Food	249	40113
Processing	2688	210685

Table 8 Robustness checks for different distance cutoffs (dc) in km for conley standard errors are presented (Corresponding to Table 2). Standard errors and the respective p-values are given. A distance cutoff of 900 km includes all distances between firms

Coef.	$dc = 50$	$dc = 100$	$dc = 250$	$dc = 500$	$dc = 900$
Constant	0.0191***	0.0205***	0.0226***	0.0211***	0.0165***
Mining	0.0357	0.0348	0.0360	0.0278	0.0220*
Chemistry	0.0239***	0.0234***	0.0236***	0.0197***	0.0146***
Services	0.0151***	0.0149***	0.0148***	0.0142***	0.0104***
Energy	0.0216*	0.0203**	0.0207**	0.0150***	0.0103***
Vehicles	0.0287	0.0291	0.0300	0.0218	0.0149
Finance	0.0106***	0.0103***	0.0102***	0.0086***	0.0064***
Trade	0.0088***	0.0081***	0.0069***	0.0062***	0.0047***
Agriculture	0.0308***	0.0308***	0.0224***	0.0184***	0.0138***
Machinery	0.0112**	0.0120**	0.0144*	0.0118**	0.0079***
Food	0.0225***	0.0216***	0.0208***	0.0167***	0.0117***
Processing	0.0100**	0.0102**	0.0079***	0.0074***	0.0053***
<i>age</i>	0.0010***	0.0009***	0.0007***	0.0006***	0.0005***
<i>ltoas</i>	0.0021***	0.0022***	0.0024***	0.0021***	0.0015***
<i>lsa</i>	0.0209***	0.0197***	0.0131***	0.0106***	0.0082***
<i>dr</i>	0.0129	0.0126	0.0126	0.0079	0.0057
<i>far</i>	0.0138	0.0149	0.0157	0.0135	0.0103
d^{sup}	0.4600***	0.4309***	0.4297***	0.3915***	0.2919***
$I_a d^{\text{sup}}$	0.5766*	0.5608*	0.5706*	0.6064*	0.4669**
d^{cus}	0.4309***	0.4012***	0.3851***	0.3593***	0.2691***
$I_a d^{\text{cus}}$	0.5280*	0.5073*	0.5061*	0.4961*	0.3836**

Notes: Significance at the 10% level (*), at the 5% level (**), and at the 1% level (***).

Table 9 Regression results (using sector construction as the reference sector). Only companies younger than 20 years are included in the regression and their median age of about 12 years is used for the interaction

Coef./Fit		Coef.	
Constant	0.2501*** (0.0230)	processing	0.0326** (0.0142)
Mining	-0.0700 (0.0625)	<i>age</i>	-0.0544*** (0.0135)
Chemistry	0.1068*** (0.0307)	<i>ltoas</i>	-0.0140*** (0.0034)
Services	0.0958*** (0.0267)	<i>lsa</i>	0.3115*** (0.0247)
Energy	0.0569* (0.0317)	<i>dr</i>	-0.0118 (0.0173)
Vehicles	-0.0333 (0.0547)	<i>far</i>	-0.0006 (0.0182)
Finance	0.0585*** (0.0158)	d^{sup}	1.2886* (0.7178)
Trade	0.0315*** (0.0108)	$I_a d^{\text{sup}}$	-0.7056 (0.7154)
Agriculture	-0.0015 (0.0688)	d^{cus}	-1.6722** (0.6692)
Machinery	0.0403** (0.0203)	$I_a d^{\text{cus}}$	0.4693 (0.5963)
Food	0.0196 (0.0374)		
Obs.	9 121		
R ²	0.0308		
Adj. R ²	0.0287		
Wald (F)	14.4711***		

Notes: Conley standard errors estimated with a distance cutoff of 250 km are given in parentheses.

* Significance at the 10% level.

** Significance at the 5% level.

*** Significance at the 1% level.

Table 10 Robustness checks for different distance cutoffs (dc) in km for conley standard errors are presented (corresponding to Table 9). Standard errors and the respective p-values are given. A distance cutoff of 900 km includes all distances between firms

Coef.	dc = 50	dc = 100	dc = 250	dc = 500	dc = 900
Constant	0.0282***	0.0270***	0.0230***	0.0204***	0.0174***
Mining	0.0631	0.0627	0.0625	0.0426	0.0306**
Chemistry	0.0367***	0.0329***	0.0307***	0.0237***	0.0170***
Services	0.0228***	0.0248***	0.0267***	0.0274***	0.0202***
Energy	0.0354	0.0312*	0.0317*	0.0263**	0.0184***
Vehicles	0.0566	0.0547	0.0547	0.0488	0.0364
Finance	0.0164***	0.0170***	0.0158***	0.0157***	0.0120***
Trade	0.0128**	0.0117***	0.0108***	0.0107***	0.0084***
Agriculture	0.0713	0.0767	0.0688	0.0605	0.0467
Machinery	0.0177**	0.0186**	0.0203**	0.0170**	0.0121***
Food	0.0483	0.0454	0.0374	0.0300	0.0232
Processing	0.0158**	0.0157**	0.0142**	0.0143**	0.0107***
<i>age</i>	0.0120***	0.0121***	0.0135***	0.0139***	0.0105***
<i>ltoas</i>	0.0030***	0.0030***	0.0034***	0.0032***	0.0025***
<i>lsa</i>	0.0318***	0.0302***	0.0247***	0.0236***	0.0183***
<i>dr</i>	0.0188	0.0175	0.0173	0.0126	0.0098
<i>far</i>	0.0196	0.0197	0.0182	0.0150	0.0111
d^{sup}	0.6393**	0.6167**	0.7178*	0.6997*	0.5399**
$I_a d^{\text{sup}}$	0.8179	0.7106	0.7154	0.6703	0.5284
d^{cus}	0.6249***	0.6026***	0.6692**	0.6419***	0.4883***
$I_a d^{\text{cus}}$	0.7788	0.6681	0.5963	0.5217	0.4163

Notes: Significance at the 10% level (*), at the 5% level (**), and at the 1% level (***).

Funding Open Access funding enabled and organized by Projekt DEAL. We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Albrecht WS, Stice EK, Stice JD (2007) Financial accounting, 1st edn. Cengage Learning
- Aleksandrova E, Behrens K, Kuznetsova M (2020) Manufacturing (co)agglomeration in a transition country: Evidence from russia. *J Regional Sci* 60(1):88–128
- Audretsch DB, Dohse D (2007) Location: a neglected determinant of firm growth. *Rev World Econ* 143(1):79–107

- Becchetti L, Trovato G (2002) The determinants of growth for small and medium sized firms. the role of the availability of external finance. *Small Bus Econ* 19(4):291–306
- Behr A, Rohwer G (2019) *Wirtschafts- und Bevölkerungsstatistik*. UVK, München
- Boeri T, Cramer U (1992) Employment growth, incumbents and entrants: evidence from germany. *Int J Ind Organ* 10(4):545–565
- de Bok M, van Oort F (2011) Agglomeration economies, accessibility, and the spatial choice behavior of relocating firms. *J Transp Land Use* 4(1):5–24
- Burger MJ, van Oort FG, van der Knaap B (2010) A treatise on the geographical scale of agglomeration externalities and the modifiable areal unit problem. *Sci Reg* 9(1):19–40
- Coad A (2007a) A closer look at serial growth rate correlation. *Rev Ind Organ* 31(1):69–82
- Coad A (2007b) Firm growth: A survey. *Doc Trav Cent Econ Sorbonne* 24:1–72
- Combes P-P (2000) Economic structure and local growth: France, 1984–1993. *J Urban Econ* 47(3):329–355
- Combes P-P, Gobillon L (2015) The empirics of agglomeration economies. In: *Handbook of regional and urban economics*, vol 5. Elsevier, Amsterdam, pp 247–348
- Dunne T, Roberts MJ, Samuelson L (1989) The growth and failure of us manufacturing plants. *Q J Econ* 104(4):671–698
- Durantón G, Overman HG (2005) Testing for localization using micro-geographic data. *Review of Economic Studies* 72(4):1077–1106
- Durantón G, Overman HG (2008) Exploring the detailed location patterns of uk manufacturing industries using microgeographic data. *J Regional Sci* 48(1):213–243
- Durantón G, Puga D (2004) Micro-foundations of urban agglomeration economies. In: *Handbook of regional and urban economics*, vol 4. Elsevier, Amsterdam, pp 2063–2117
- Ellison G, Glaeser EL, Kerr WR (2010) What causes industry agglomeration? evidence from coagglomeration patterns. *Am Econ Rev* 100(3):1195–1213
- Eurostat (2008) *Nace rev. 2—statistical classification of economic activities in the european community. Methodologies and working papers*. General and regional statistics, Theme
- Fagiolo G, Luzzi A (2006) Do liquidity constraints matter in explaining firm size and growth? some evidence from the italian manufacturing industry. *Ind Corp Change* 15(1):1–39
- Feser EJ (2002) Tracing the sources of local external economies. *Urban Stud* 39(13):2485–2506
- Fuertes-Callen Y, Cuellar-Fernandez B (2019) Inter-relationship between firm growth and profitability in a context of economic crisis. *J Bus Econ Manag* 20(1):86–106
- Glaeser EL, Kerr WR (2009) Local industrial conditions and entrepreneurship: how much of the spatial distribution can we explain? *J Econ Manag Strategy* 18(3):623–663
- Glaeser EL, Kallal HD, Scheinkman JA, Shleifer A (1992) Growth in cities. *J Polit Econ* 100(6):1126–1152
- Goldman B, Klier T, Walstrum T (2019) Within-industry agglomeration of occupations: Evidence from census microdata. *J Regional Sci* 59(5):910–930
- Greenstone M, Hornbeck R, Moretti E (2010) Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *J Polit Econ* 118(3):536–598
- Harhoff D, Stahl K, Woywode M (1998) Legal form, growth and exit of west german firms—Empirical results for manufacturing, construction, trade and service industries. *J Ind Econ* 46(4):453–488
- Henderson V, Kuncoro A, Turner M (1995) Industrial development in cities. *J Polit Econ* 103(5):1126–1090
- Henrekson M, Johansson D (2010) Gazelles as job creators: a survey and interpretation of the evidence. *Small Bus Econ* 35(2):227–244
- Hoogstra GJ, van Dijk J (2004) Explaining firm employment growth: does location matter? *Small Bus Econ* 22(3):179–192
- Illy A, Hornych C, Schwartz M, Rosenfeld MTW (2009) Urban growth in germany – the impact of localization and urbanization economies. *IWH discussion papers*, vol 19/2009
- Jofre-Monseny J, Marín-López R, Viladecans-Marsal E (2011) The mechanisms of agglomeration: Evidence from the effect of inter-industry relations on the location of new firms. *J Urban Econ* 70(2-3):61–74
- Klier T, McMillen DP (2008) Evolving agglomeration in the u.s. auto supplier industry. *J Regional Sci* 48(1):245–267
- Kolko J (2010) Urbanization, agglomeration, and coagglomeration of service industries. In: *Agglomeration economics*. University of Chicago Press, pp 151–180
- Levratto N, Zouikri M, Tessier L (2010) The determinants of growth for smes—a longitudinal study from french manufacturing firms. *CNRS-EconomiX (Technical report)*
- Lopez-Garcia P, Puente S (2012) What makes a high-growth firm? a dynamic probit analysis using spanish firm-level data. *Small Bus Econ* 39(4):1029–1041
- Marshall A (1920) *Principles of economics* macmillan, 8th edn. (1890)

- Mathias BD, McCann BT, Whitman DS (2021) A meta-analysis of agglomeration and venture performance: Firm-level evidence. *Strateg Entrepreneursh J* 15:430–453. <https://doi.org/10.1002/sej.1390>
- McCann BT, Folta TB (2009) Demand-and supply-side agglomerations: distinguishing between fundamentally different manifestations of geographic concentration. *J Management Studs* 46(3):362–392
- Nakamura R, Paul CJM (2019) Measuring agglomeration. In: *Handbook of regional growth and development theories*. Edward Elgar,
- van Oort FG (2007) Spatial and sectoral composition effects of agglomeration economies in the Netherlands. *Pap Reg Sci* 86(1):5–30
- Penner SJ (2004) *Introduction to health care economics & financial management: fundamental concepts with practical applications*, 1st edn. Lippincott Williams & Wilkins, Philadelphia, Baltimore, New York, London
- Regionaldatenbank Deutschland (2020) Erwerbstätigenrechnung d.Bundes u.d. Länder. <https://www.regionalstatistik.de> (Code: 13312). Accessed: 21.12.2020
- Rigby DL, Essletzbichler J (2002) Agglomeration economies and productivity differences in us cities. *J Econ Geogr* 2(4):407–432
- Rosenthal SS, Strange WC (2001) The determinants of agglomeration. *J Urban Econ* 50(2):191–229
- Rosenthal SS, Strange WC (2004) Evidence on the nature and sources of agglomeration economies. In: *Handbook of regional and urban economics*, vol 4. Elsevier, Amsterdam, pp 2119–2171
- Schneider O, Lindner A (2010) The value of lead logistics services. In: Vallespir B, Alix T (eds) *Advances in production management systems: new challenges, new approaches*, pp 315–322
- Schreyer P (2000) High-growth firms and employment. OECD science, technology and industry working papers
- Statistisches Bundesamt (2008) *Klassifikation der Wirtschaftszweige*. Statistisches Bundesamt, Wiesbaden
- Statistisches Bundesamt (2017) *Volkswirtschaftliche Gesamtrechnungen – Input-Output-Rechnung – 2013 (Revision 2014, Stand: August 2016) – Fachserie 18 Reihe 2*
- Wand MP, Jones MC (1995) *Kernel smoothing*. Monographs on statistics and applied probability, vol 60. Chapman & Hall, London

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.