Estimating the Skill Bias in Agglomeration Externalities and Social Returns to Education

Groot, Stefan P.T.; de Groot, Henri L.F.

published in
Economist (Netherlands)
2020

DOI (link to publisher)
10.1007/s10645-019-09354-w

document version
Publisher's PDF, also known as Version of record

De Economist (2020) 168:53–78
https://doi.org/10.1007/s10645-019-09354-w

Estimating the Skill Bias in Agglomeration Externalities and Social Returns to Education: Evidence from Dutch Matched Worker-Firm Micro-Data

Stefan P. T. Groot[1] · **Henri L. F. de Groot**[2,3,4]

Published online: 26 November 2019

Abstract

This paper employs a unique set of micro-data covering almost one-third of the Dutch labor force, to estimate the heterogeneity of agglomeration externalities across education levels. This paper shows that there is substantial heterogeneity in the relationship between agglomeration and productivity of workers (proxied by their hourly wage) with different educational background. Apart from estimating the impact of the aggregate density of regional labor markets, we also estimate whether the composition of the local labor market in terms of education is related to the productivity of different types of workers. Using the presence of universities as an instrument, we estimate the effect of the supply of university graduates on wages, i.e. the social return to education. We find that agglomeration externalities are substantially higher for high- and medium skilled than for low-skilled employees. We find no positive effects from the presence of high-skilled on the productivity of low-skilled.

Keywords Agglomeration, Education, Knowledge-spillovers, Wages, Local labor markets

JEL Classification J3 · I2

During the first stage of working on this paper, Stefan Groot was working at the Vrije Universiteit Amsterdam. We are grateful to two anonymous reviewers of this journal for their constructive comments on an earlier version of this paper. We are also grateful to Coen Teulings and Wouter Vermeulen for many inspiring discussions on this and related papers. The usual disclaimer applies.

✉ Stefan P. T. Groot
s.p.t.groot@cpb.nl

1 CPB Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands

2 Department of Spatial Economics, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

3 Tinbergen Institute, Amsterdam, Rotterdam, The Netherlands

4 Ecorys NEI, Rotterdam, The Netherlands

In almost all countries, there is a constant migration towards the towns.

The large towns […] absorb the very best blood from all the rest […].
the most enterprising, the most highly gifted, those with the highest physique.
and the strongest characters go there to find scope for their abilities.
Alfred Marshall (1890).

1 Introduction

As the often cited quote of Alfred Marshall illustrates, cities (or more generally: agglomerations that combine large economic mass and high density) are assumed to work like powerful magnets attracting the most skilled and able employees from the surrounding areas. Agglomerated areas in part have an above-average wage level because of their more favorable labor market composition, consisting of more high-educated and specialized employees. But additionally, it is a well-known fact that even after correcting for regional differences in labor market composition, agglomeration is associated with higher levels of productivity and wages. The debate in the literature that addresses agglomeration externalities is not so much about the question whether such externalities exist, but it is mostly about their extent. The main focus in the literature is on how to properly estimate the size of agglomeration externalities, and how to identify the mechanisms that drive those agglomeration externalities.

Since many theories that contribute to our understanding of agglomeration externalities are either related to specialization [following the work of Smith (1776)], or to knowledge spillovers [building on the seminal work of Marshall (1890) and Jacobs (1969)], there is likely to be a very strong interdependency between the level of education and the extent of agglomeration externalities. The reasons for this are straightforward: high-educated individuals are generally more specialized, are in the possession of more knowledge, and are more likely to perform tasks that are related to processing knowledge, information, and innovation. Combined with the tacit nature of knowledge, one can indeed expect the returns from density to be particularly high for high-skilled. The first research question that is addressed in this paper is therefore to what extent the relationship between agglomeration—which we define as the density of jobs in the local labor market—and productivity as proxied by hourly wages is different for workers with different levels of education.

The estimation of agglomeration economies is very sensitive to the geographic units that are used, and the use of administrative boundaries is likely to introduce arbitrary border effects (Briant et al. 2010; Combes and Gobillon 2015). Therefore, we base our definition of local labor markets on a distance decay function that is fitted on the data. Using hourly wages as a proxy for productivity is a common approach in the literature (cf. Glaeser and Maré 2001; Combes et al. 2008; Combes and Gobillon 2015). For workers, higher wages might be offset by higher housing costs or commuting costs, which ensures spatial equilibrium. From the perspective of the firm, however, choosing a location with high labor costs and high land prices will only be justified if it offers a productivity advantage (Puga 2010; Groot et al. 2014).

The second question that is addressed in this paper is whether the presence of a high-skilled workforce in regions with a high economic density may *itself* explain part of the wage differential with less dense areas. The knowledge spillovers discussed by Marshall (1890) and Jacobs (1969) are not only more likely to occur in a *denser* environment because there are more interactions where knowledge can be exchanged, but they are also more likely to occur in a *more knowledge intensive* environment in which there is a higher regional stock of knowledge that can be exchanged. As a long strand of literature argues (see, for example, Moretti 2004), the local stock of knowledge has characteristics of a local public good which has a positive impact on the wages of both high- and low-educated workers, thereby generating substantial social returns to education.

There can also be social returns to education for other reasons than productivity, for example due to the correlation between the level of education and social outcomes such as crime, but these types of social returns are not the topic of this paper. Our definition of the social return is—similar to Acemoglu and Angrist (2000) and Moretti (2004)—the effect of an increase of the share of high-educated workers in a local labor market on total wages *minus* the private returns to education which are measured as the *ceteris paribus* wage differential between workers with different levels of education. Rather than having a positive effect on local wages, increased supply of high-educated workers could also *reduce* the average wages in a region. Because of the law of supply and demand, more supply of high-educated workers at a given demand could simply reduce wages of high-skilled workers, and also the wages of employees with lower levels of education can be negatively affected due to replacement effects of low-skilled by high-skilled.

Because there is a substantial overlap—and potential interdependence—between the effects of high levels of agglomeration and high levels of education, it is important to separate both forces in an integrated empirical framework. As both the estimation of agglomeration externalities and the estimation of the social returns to education are affected by issues related to endogeneity and unobserved heterogeneity, we employ a mix of empirical techniques in an effort to avoid some common pitfalls. To avoid biased estimates due to endogeneity of agglomeration, we use pre-industrial revolution density as an instrument for current density of local labor markets. The use of long lags of population density as an instrument for current density is a standard practice since the work of Ciccone and Hall (1996) (cf. Combes et al. 2010; Combes and Gobillon 2015). As an instrument for the share of high-educated workers, we use the local supply of university graduates (an approach similar to that of Moretti 2004). The availability of an extensive set of microdata covering over two million employees and an entire decade (2000–2010) allows us to control for individual worker characteristics and industry effects. However, to exclude the possibility that unobserved heterogeneity still affects the estimation results, we repeat our estimates whilst including individual worker fixed effects, thereby exploiting the panel structure of our dataset.

From a policy perspective, the questions addressed in this paper are relevant for several reasons. Policy decisions related to the housing market (e.g. considering where to build and for whom to build), investments in infrastructure (which may result in shifts in employment because the costs of commuting to a more productive location may have changed), as well as decisions related to the location of institutions of education may all affect both economic density and the share of high-educated workers in the local labor market. This in turn may not only have consequences for those who are directly involved, but it may also have substantial consequences for all other employees in the local labor market, which may vary for workers with different levels of education. Also, demographic trends such as population decline in peripheral regions are often skill biased. For example, between 2000 and 2010, the share of high-educated employees in South Limburg (which is one of the regions in the Netherlands experiencing a population decline) increased by 5.1 percent points while it increased by 7.7 percentage points on average in the Netherlands.[1] If the presence of high-skilled employees increases the productivity of low-skilled employees, this may not only result in increased regional wage inequality, but lagging wage and productivity growth may also contribute to further population decline.

The remainder of this paper is structured as follows. The next section will provide a discussion of different theories that potentially explain the relation between agglomeration and productivity, paying special attention to the importance of education and knowledge spillovers. Also, it provides an overview of the relevant empirical literature. Section 3 describes the different data sources that are used, and will present a number of stylized facts. Section 4 discusses our empirical strategy. Estimation results of different specifications are presented in Sect. 5. We compare results estimated using OLS to estimates using IV, and also present results that include worker fixed effects. Section 6 concludes.

2 Theoretical and Empirical Literature

2.1 Agglomeration Economies

Wages vary across regions for several reasons. Combes et al. (2008) distinguish between three main sources of regional wage disparities: (1) composition of the local labor market, (2) the availability of local non-human endowments that increase productivity (such as access to the sea), and (3) agglomeration economies which are the main topic of this paper. Agglomeration economies are productivity differences that follow *ceteris paribus* from close proximity between different firms and consumers, thick labor markets, and from knowledge spillovers. As highlighted in the introduction, the existence of such a positive relationship between agglomeration and labor productivity is a well-known stylized fact. Much less, however, is known about the mechanisms that could cause the relationship between agglomeration and productivity. Many different—sometimes even opposing—theories have been developed that contribute to our understanding of agglomeration externalities. Perhaps

[1] Source: own calculations based on data from CBS Statline (http://statline.cbs.nl).

one of the most important is that cities allow for more specialization because of their larger market size. As already noted by Adam Smith (1776), increased specialization results in higher productivity—and therefore in higher wages and wealth. As a narrower set of tasks to be executed will result in a higher ability to perform those tasks, it reduces the costs of switching between different tasks. Also, it makes the application of technology less complicated.

In more general terms, the existence of firm-level increasing returns to scale in combination with non-tradable products and transportation costs is known to result in regional differences in productivity (Henderson 1988; Fujita 1989; Ciccone and Hall 1996). The importance of specialization, non-tradables, and transportation costs also points at a potential cause for differences in agglomeration externalities for workers with different educational background. The work of high-educated employees is relatively specialized and complex compared to the tasks typically performed by low-educated employees. Their work is also more likely to be related to processing information, thus involving more tacit knowledge and requiring more face-to-face contact, making the tasks they perform less tradable. Even within the manufacturing sector, high-skilled workers often perform work that is not directly related to the production process. Therefore, it is likely that the relationship between agglomeration and productivity is stronger for high-educated workers.

Following Marshall (1890), several other explanations for the existence of agglomeration economies have been proposed. They may arise because of the thick labor markets that come with agglomeration, due to linkages between firms, or because of knowledge spillovers (for an overview of this literature, see Rosenthal and Strange 2004, or Duranton and Puga 2004, who provide an extensive overview of the microeconomic foundations of agglomeration economies). The large size of local markets results in lower transaction costs on markets for intermediaries and final goods (Harvey 1981), lower transaction costs on the labor market as well as a higher probability that a good match is established between employers and employees, and it reduces the costs of incomplete information (Duranton and Puga 2004).

Close proximity also facilitates the exchange of knowledge, resulting in more innovation (Jaffe et al. 1993). Externalities that take place between industries are in the agglomeration literature generally referred to as urbanization economies, while externalities that take place within industries are referred to as localization economies (Fujita et al. 1999; Fujita and Thisse 2002). Glaeser et al. (1992) analyze different types of knowledge spillovers. Marshall-Arrow-Romer externalities (as Glaeser et al. label intra-sectoral knowledge spillovers) assume that knowledge is industry specific. In their view, exchanging knowledge takes place mostly when firms with similar activities are in close proximity. In contrast, Jane Jacobs (1969) argues that knowledge spillovers are more likely to occur *between* rather than *within* industries because the larger differences between industries provide more opportunities for learning. While agglomeration economies are generally thought to arise when concentration is high (because this implies a specialized regional economy with high returns to scale), Porter (1990) argues that it may be better if there are many competing firms within industries in a region because high competition forces firms to increase productivity. Even though Glaeser et al. (1992) found that Jacobs externalities are generally the most important from an empirical point of view, later reviews of the literature have shown that results regarding the importance of different types of agglomeration economies are in fact very mixed (Rosenthal and Strange 2004; Beaudry and Schiffauerova 2009; Melo et al. 2009; De Groot et al. 2016).

Because of the simple fact that high-educated employees have accumulated a larger knowledge base and because their work is much more likely to involve handling information, knowledge, complexity, or creativity rather than production work, it is very likely that the probability that knowledge-spillovers will occur is larger for high-educated workers. Consequently, this should result in a stronger relationship between agglomeration and wages for high-educated workers. The presence of high-educated workers itself may in that case be one of the driving forces behind agglomeration economies. Also, if knowledge spillovers are indeed an important cause of agglomeration economies, it is possible that it is not (or at least not only) agglomeration in general that matters, but particularly agglomeration of high-educated workers.

2.2 Social Returns to Education

Following the work of Schulz (1988) and Rauch (1993), the presence of high-skilled employees in a region is often considered as a local public good. There are several ways through which knowledge spillovers could result in a higher productivity. For example, they may provide the firm with knowledge about new technologies that increase productivity, they may transfer parts of their knowledge to other employees who become more productive as a consequence, or they may be complementary to the knowledge of different types of workers. What these mechanisms have in common is that they result from "the sharing of knowledge and skills between workers that occurs through both formal and informal interaction" (Rauch 1993, p. 380).

As Rauch (1993) notes, the work of Jacobs (1969) provides numerous examples of the ways through which interaction between educated and skilled individuals can have a positive effect on productivity. Although part of the knowledge spillovers taking place in cities may be considered part of agglomeration externalities in general—e.g. because higher density increases the potential for the interaction and the exchange of knowledge—there is also a part that can be attributed to the local knowledge stock. If we would vary the share of high-skilled workers in the local labor market at a given economic density, it is likely that the extent of knowledge spillovers will change substantially. We consider this to be the effect of the social returns to education rather than the effect of agglomeration.

Lucas (1988) models the stock of human capital as a Hicks neutral shift in technology, resulting in a shift in the production function that allows for a higher level of productivity of all other inputs. It is, however, likely that its effect will differ across different production factors.

2.3 Empirical Literature

A growing body of literature provides us with empirical estimates of either agglomeration externalities or knowledge spillovers. The interaction between the two has attracted little attention until now. Traditionally, agglomeration externalities have often been estimated on regional level data, using average wages or value added as productivity measures (for example, Ciccone and Hall 1996). Because cross-country micro-data are relatively scarce, this holds in particular for international comparative studies such as Ciccone (2000). More recently it has been argued that the use of aggregated data fails to sufficiently control for worker and firm heterogeneity (which are, as the previous subsection discussed, itself important sources of regional wage differences), resulting in an upward bias of estimated agglomeration externalities (Combes et al. 2008; Duranton 2010; Puga 2010; Groot et al. 2014). Indeed, as Melo et al. (2009) show in their meta-study, the use of aggregate data tends to result in relatively high estimated agglomeration economies. The use of micro-data is thus essential to address complex spatial-economic questions (see also Van Bergeijk et al. 2011). Studies that are based on micro-data often rely on augmented Mincerian wage regressions that use wages as a proxy for the productivity level of individual workers (Glaeser and Maré 2001; Combes et al. 2008; Groot et al. 2014; Verstraten et al. 2019). Melo et al. (2009) find that agglomeration elasticities are generally estimated to be in the 3–8% range, whereby studies using micro-data are generally in the lower half of that range. Using Dutch micro-data, Groot et al. (2014) estimate that doubling the economic density in Dutch NUTS-3 regions is associated with a 4.8% increase in wages of employees with given observed individual characteristics working in the same sector. Besides using data on wages, there is also a smaller literature estimating agglomeration externalities from firm level productivity (for example, Henderson 2003). Groot and Weterings (2013) estimate the relationship between employment density and firm-level total factor productivity (TFP) for Dutch NUTS-3 regions, and find no evidence for agglomeration externalities based on TFP. They conclude that higher productivity is offset by higher wages and land rents.

Following the work of Rauch (1993), who finds a positive effect of the average level of education in a US metropolitan area, several attempts have been made to estimate the social returns to education, some of which take interaction effects with the level of education into account. Moretti (2004), who also uses US data, finds evidence for relatively large benefits of the presence of high-skilled workers for low-skilled workers. He finds that working in an area with a relatively high share of high-educated workers is beneficial for high-skilled workers as well, but the effect is smaller. Canton (2009)—who uses Dutch data—finds some limited evidence for knowledge spillovers resulting from the presence of high-educated workers, but finds it to be limited to spillovers within firms.

3 Data and Stylized Facts

3.1 Data Sources

This paper relies on a number of rather unique micro datasets that have been made available by Statistics Netherlands (CBS). They cover the years 2000–2010, 2.1 million (anonymous) employees, and 11.5 million observations. The available datasets—each containing different types of characteristics related to employees or firms—have been merged into one large file for our analyses.

At the basis is a large fiscal dataset that includes employer reported pre-tax wages and hours worked of all employees in the Netherlands (*Polisadministratie*). Our wage definition includes all monetary regular and incidental payments (the latter include performance payments, vacation payments, if applicable the 13th month, and payments for overwork, amongst other things), and also the monetary value of payments in kind (such as the use of a company-provided car). We have excluded irregular bonuses and golden handshakes, because these may cover rewards for multiple years. Wages have been deflated using the Price Deflator (CPI, *Consumenten Prijs Index*) of Statistics Netherlands. This dataset has been merged to census data (SSB, *Sociaal Statistisch Bestand*), which includes several individual characteristics such as age, gender, country of birth, and all past and present addresses that are collected in a database maintained by municipalities (GBA, *Gemeentelijke Basis Administratie*). With respect to country of birth, we distinguish three different categories: native, OECD, and non-OECD.[2] For information on firms (in particular sector), we rely on the firm registry of Statistics Netherlands (ABR, *Algemeen Bedrijfs Register*).

As the three aforementioned datasets are exclusively taken form registries, the resulting dataset includes all Dutch employees that are required to pay taxes in the Netherlands and have a current address in the Netherlands (it, however, excludes cross-border commuters and self-employed). For the work location of employees, we rely on two different data sources. First: for employees that work for firms that have all their activities in only one known municipality in the Netherlands—as known through the regionalized version of the firm registry (*ABR Regiobase*)—we take that location. For firms with establishments in multiple regions, we use the most likely work location as determined by Statistics Netherlands in the municipality of work registry (*Gemeente Standplaats*).[3] Additionally, we add (time invariant) data on level and type of education from the Dutch education registry (*Opleidingenregister*). In this registry, data from several sources such as the labor force survey (EBB, *Enquête Beroepsbevolking*) and different diploma registries have been collected. This step results in a substantial reduction of our sample size, as more than half of Dutch employees are not included in this dataset. There is a considerable underrepresentation of old workers in the education registry. This may result in a downward bias in our estimates of agglomeration economies, as the literature has shown that agglomeration economies are larger for more experienced workers (De la Roca and Puga 2017). Also, higher levels of education were included in the registries from an earlier point in time. However, this does not affect our results as we estimate only on split samples.

[2] We have excluded employees born in Turkey from the OECD group, and added them to the non-OECD group.

[3] CBS derives local employment by combining tax data (that give total employment per firm) with a survey where multi-establishment firms with 10 or more employees provide employment in each municipality. Employees of multi-establishments firms with less than 10 employees (with a relatively low share in employment), are allocated to the headquarter.

In most of our analyses we distinguish four different levels of education: low, medium, college, and university graduates. Low-educated employees are defined as individuals that have received at most a VMBO or MBO 1, 2 or 3 diploma (these are the lower types of secondary and tertiary education, with a generally practical focus), medium educated as employees with a HAVO or VWO diploma (which are the highest levels of Dutch secondary education, with a theoretical orientation and focus on later enrolment in higher tertiary education) or MBO 4 diploma (an intermediate level of tertiary education with a generally theoretical orientation), while we define college graduates as individuals with a HBO (positioned just below the level of a university) diploma or a university BSc degree. Employees with an MSc or a Phd degree are classified as university graduates.

Several selection criteria have been applied. Because employees who work through employment or pay-roll agencies are registered at the municipality where these agencies are located, they have been excluded from our analyses (as their actual work location is unknown). The level of observation in our analyses is that of the job (an employee can have multiple jobs during a year). We take only the highest-paid job of each individual employee in each year into account. We have removed all jobs with a duration of less than 1 month or less than 12 h per week, and those earning less than the minimum wage. Also, only employees between 18 and 65 have been included. Agriculture and the public sector have been excluded from our analyses, because wages in these sectors are to a lesser extent affected by regional forces. In addition to this, public sector employees are not well represented in the registries that we use to determine work location.

3.2 Descriptive Statistics

Table 1 presents a number of descriptive statistics about our data. We present separate descriptives by level of education, but have pooled the 11 cross-sections that are available. On average, an individual employee is observed 5.1 (lower educated) to 5.7 times (college graduates) over the period of 11 years, emphasizing the panel structure of our data (although it is clearly not a balanced panel). As expected, a higher level of education corresponds to substantially higher hourly wages. Also, wages of high-educated employees show more variation. A likely explanation for the latter is that wages of low-educated employees are to a larger extent institutionalized and downwards constrained by minimum wages. High-educated workers are on average younger than low-educated (most likely due to the fact that younger cohorts are more highly educated), more likely to work part-time, and more likely to have been born in the Netherlands—which holds in particular for employees that were born in non-OECD countries. Descriptive statistics related to age and country of birth have to be interpreted with caution, however, as both older and foreign born employees are underrepresented in the education registry. These figures apply thus only to our sample and are not representative to the situation on the Dutch labor market.

Table 1 also includes two key variables related to the local labor market where each individual lives. The employment density is the total number of jobs in the

Table 1 Descriptive statistics by level of education, 2000–2010

Dep.: log hourly wage	Low	Medium	College	University
#Observations	3,755,289	3,225,783	2,877,818	1,641,720
#Employees	736,962	585,415	508,998	291,412
Hourly wage (EUR)	15.01 (5.94)	18.99 (15.85)	22.26 (12.32)	28.95 (23.36)
Log hourly wage	2.659 (0.30)	2.847 (0.41)	3.013 (0.406)	3.234 (0.481)
Age (years)	39.87 (11.31)	36.16 (9.97)	33.95 (8.35)	34.96 (7.83)
Female (share)	0.296	0.319	0.307	0.334
Part-time (share)	0.246	0.234	0.203	0.196
Foreign born, OECD (share)	0.018	0.018	0.017	0.029
Foreign born, non-OECD (share)	0.089	0.039	0.032	0.039
Employment density (jobs local labor market)	317,481 (199,373)	348,660 (209,355)	370,979 (208,951)	447,872 (207,740)
Share of college and university graduates in the local labor market	0.433 (0.048)	0.441 (0.048)	0.447 (0.047)	0.464 (0.044)

Standard deviations are in parentheses

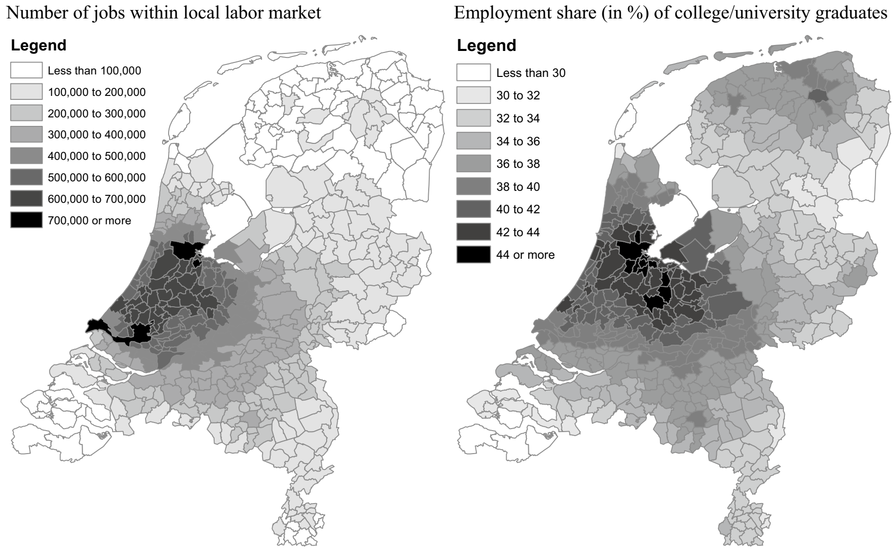

Fig. 1 Agglomeration and the share of high-educated employees

relevant local labor market around each individual, which has been aggregated directly from the micro-data (recall that we have employment available for almost all Dutch jobs excluding those working for employment and payroll agencies). Rather than using an administrative regional classification, we apply a dynamic classification that takes each municipality in the Netherlands as the center of the relevant local labor market for that municipality. The extent to which surrounding municipalities are considered to be part of the local labor market around each municipality depends on a distance decay function, which has been estimated on commuter flows. "Appendix 1" elaborates on our classification of local labor markets and the distance decay parameter estimation process.

Because the effective size of each local labor market is the same (e.g., the distance decay function that was used is time and space invariant), the total number of jobs within it actually measures density. High-educated employees are substantially more likely to live in a local labor market with a high employment density. Also, they are more likely to live in an area with a high share of high-educated employees, although the difference with workers with other types of education is not large.

3.3 Stylized Facts

Figure 1 presents the effective employment density for the local labor market centered around each municipality in the Netherlands (left panel) as well as the share of highly educated employees within that local labor market (right panel). All figures are related to the location where an individual *works* rather than where he *lives*. Effective density ranges from far less than 100 thousand jobs in a number of peripheral regions, to more than 700 thousand in Amsterdam and Rotterdam. Even though there is a number of large cities (such as Groningen) outside the Randstad region (the area in the West of the country where most large agglomerations are located), effective density remains relatively low because these agglomerations are surrounded by a hinterland with a relatively low density. As the right panel of Fig. 1 shows, there is a strong positive relationship between employment density in a region and the local share of higher educated employees, although Rotterdam—with a relatively low-educated workforce—is a notable exception. In contrast, a number of cities in peripheral regions—most notably Groningen, which has a large university—have a relatively high share of high-educated jobs as well.

Figure 2 presents average hourly wages for employees conditional on their level of education. When we compare the three panels, an interesting pattern emerges. While the average wage of low-educated employees in local labor markets has a range of only 0.60 euro (e.g. from 14.10 to 14.70), this is about 2 euro for medium educated, 1.50 for college graduates and 3.00 for university graduates. The scale in the different maps is of course somewhat subjective (since it was manually chosen), but the general pattern is clear. The wages of both medium- and high-educated employees show much more variation across space compared to low-educated employees. An additional difference between low-educated employees and both medium- and high-educated employees is that the correlation between average wages and employment density that was shown in Fig. 1 is not as strong: wages of low-educated are relatively high along the entire coast line, even in the relatively peripheral region of Zeeuws-Vlaanderen (close to the border with Belgium). Even though there is a substantial difference in the level of wages, the distribution of average wages across space of medium educated employees looks remarkably similar to that of high-educated employees. The next sections will investigate to what extent the raw patterns that can be observed from the figures remain if we control for worker heterogeneity, and will further explore the driving forces behind these patterns using regression analyses.

4 Methodology

4.1 Estimation Strategy

Similar to many other studies in the literature (e.g., Moretti 2004; Combes et al. 2008; Groot et al. 2014), our empirical strategy revolves around the estimation of augmented Mincerian wage regressions (cf. Mincer 1974). In these regressions, individual wages are explained by a set of individual worker characteristics, as well as log employment density and the log share of high-educated employees within the regional labor market. Because there is likely to be substantial heterogeneity in the relationship between different worker characteristics, we estimate separate regressions for each level of education. For example, it is possible that the male–female wage differential or the effect of experience (proxied by age)

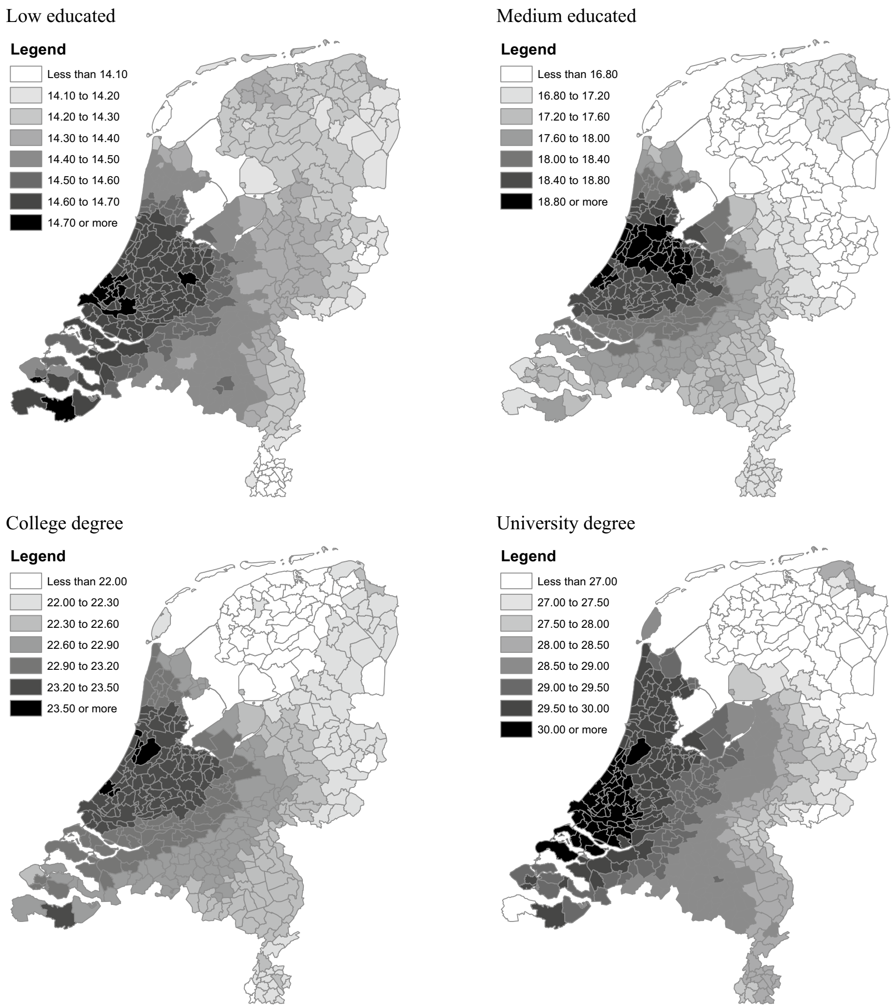

Fig. 2 Average hourly pre-tax wages by level of education

varies across workers with different levels of education. Because there may still be substantial heterogeneity in the level of education within each of the three main educational groups, we include between one and three dummies to control for sub-classifications of education levels within the four main categories in each of the regressions. To control for sectoral heterogeneity we include industry dummies on the 2-digit NACE rev. 1.1 level (in the Netherlands known as the SBI-1993 classification). Furthermore, we include a set of year dummies. Formally, our regression equation can be described as follows:

$$\begin{aligned}\log(w_{i,t}) = \alpha &+ \beta_1 age_{i,t} + \beta_2 age_{i,t}^2 + \beta_3 D_i^{female} + \beta_4 D_i^{OECD} + \beta_5 D_i^{non\text{-}OECD} \\ &+ \beta_6 D_{i,t}^{part\text{-}time} + \sum_{edu} \beta_{7,edu} D_i^{edu} + \sum_{ind} \beta_{8,ind} D_{i,t}^{ind} + \sum_{year} \beta_{9,year} D_{i,t}^{year} \\ &+ \beta_{10} \log(density_{i,t}) + \beta_{11} \log(share\ highly\ edu_{i,t}) + \varepsilon_{i,t},\end{aligned} \tag{1}$$

whereby $w_{i,t}$ is the natural logarithm of the pre-tax individual hourly wage of worker i in year t, D^* represent dummy's for different variables, and $\varepsilon_{i,t}$ is an error term that varies by individual worker and year. The reason to include the logarithms of density and the share of high-educated workers is that the resulting estimates can be interpreted as elasticities.

A difference between our approach and that of Combes et al. (2008), Groot et al. (2014) and Verstraten et al. (2019) is that they use a two-stage estimation approach, while we have estimated both individual and regional level variables at once. Even though we share the critique expressed by Combes et al. (2008)—which is derived from Moulton (1990) who shows that estimating the effects of aggregated variables (such as agglomeration and level of education) on micro units will result in a downward bias in standard errors of the estimates—the calculation of robust standard errors (clustered on the level of regions) seems a more straightforward solution for this issue.

4.2 Accounting for Endogeneity

A common problem in the literature is the endogeneity of both agglomeration and the average level of education. The cause of this is simple: agglomeration does not only increase productivity, but high productivity does also attract new workers thus resulting in higher agglomeration. To account for this endogeneity, we instrument employment density in 2000–2010 with density in 1840 which is an approach similar to the one used by Ciccone and Hall (1996), Combes et al. (2008) and Graham et al. (2010). Because historical population is not a result of current productivity, while at the same time current and historical densities are highly correlated, it is suitable as an instrument. A potential pitfall would occur if current and historical productivity are both caused by the same forces, which affected both current and historical agglomeration in a similar way. In that case, the instrument would not be sufficiently exogenous. To avoid this, we use population density prior to the industrial revolution, assuming that the present economic structure and the driving forces behind agglomeration have changed over the course of time.

As noted by Moretti (2004) and Canton (2009), the estimation of the social returns to education involves endogeneity issues as well. The reason for this is similar: wages in a region may not only be high because the share of high-educated workers is relatively high (at a given level of agglomeration), but it may also be the case that local composition of the labor market in terms of education simply reflects regional variation in the relative productivity of different types of labor. To instrument for the share of high-educated workers, we use a somewhat similar—although not the same—approach as Moretti (2004) who uses the presence of universities built under an historical government program as an instrument. Even though graduates are free to choose their work location after graduation, the presence of frictions makes the presence of universities a suitable instrument.

Rather than using a dummy variable to indicate whether a university is present, we have merged students that entered the labor market between 1995 and 2010 to the university from which they are most likely to have graduated. Because we do not know the actual institution from which an individual graduated and the graduation year, we assign graduates to the institution that offers their level of education that is closest to their residence municipality on May 1 of the year in which a student with that level should have nominally graduated. If the distance between this municipality and the nearest institution is more than 15 km, they have been excluded (approximately 9% of Dutch graduates are dropped in this step). To check the accuracy of this linking process, we have determined what percentage of graduates with ten different rare (e.g., that are offered by at most two institutions) 4-digit SOI-2006 education types has been linked to the institutions that offer these types of education. The average performance was 89%, which implies that our approach works rather well.

We distinguish four different types of higher education: HBO, BSc, MSc and Phd. An important assumption in using the local supply of university graduates as an instrument for the share of high-educated workers in the local labor market is that prospective students do not take the wage level of the area around universities into account when choosing a university. This seems plausible, because even if prospective earnings would be relatively important to students (as opposed to intrinsic motivation), students are free to migrate to a more productive region after graduation, while at the same time the costs of housing tend to be lower in more peripheral regions with lower levels of productivity.

In Sect. 5, we present both estimates that have been obtained using OLS and IV estimates. This allows us to investigate to what extent endogeneity affects the results. If the difference between these estimates is relatively large, it means that endogeneity is indeed an issue.

4.3 Fixed Effects

Even though we have a rather rich set of control variables available that substantially limit unobserved worker heterogeneity, it is not unlikely that unobserved heterogeneity remains to result in biased estimates. In particular, this could happen when specific groups of workers with narrowly defined characteristics (for example, high-educated Dutch born males at a given age) are more likely to work in more agglomerated areas when they are more productive and earn higher wages (because they are more skilled or more ambitious) compared to their colleagues with similar characteristics that work in less dense areas. If such a change in location occurs, it is reasonable to assume that individual worker characteristics remain the same.

The inclusion of worker fixed effects, as is done by, for example, Glaeser and Maré (2001), and Combes et al. (2008, 2010) is an often used method to control for worker heterogeneity. Because of the panel structure (covering a relatively long period), our data are reasonably well suited for this estimation strategy. Because our instruments are time invariant, the effects of agglomeration and the share of education are only identified on employees who change jobs to a different region when applying instrumental variables. Therefore, we drop employees working in the same municipality in all years. The formal equation that we estimate when including worker fixed effects is as follows:

$$\log(w_{i,t}) = \alpha + \beta_1 age_{i,t} + \beta_2 age_{i,t}^2 + \beta_3 D_{i,t}^{part\text{-}time} + \sum_{year} \beta_{4,year} D_{i,t}^{year} + \beta_5 \log(density_{i,t}) + \beta_6 \log(share\ highly\ edu_{i,t}) + \delta_i + \varepsilon_{i,t}, \quad (2)$$

where by the worker fixed effects are denoted by δ_i.

Even though the inclusion of fixed effects solves some econometric problems, it also has a number of drawbacks (see, for example, Wooldridge 2002). Even though the inclusion of fixed effects fully solves the problem of time-invariant omitted variable biases, time-variant omitted variables (such as worker skills and experience) may still result in somewhat biased estimates. It is, for example, possible that workers who accumulate more knowledge and abilities throughout their careers and therefore become more productive over time, are more likely to move to agglomerated areas (cf. Verstraten et al. 2019). Another problem is that the identification of agglomeration externalities through workers that move to an employer in a different region may be prone to selection bias, because the probability that an employee accepts a job offer is likely to be related to how favorable a job offer is while at the same time job offers from a region with relatively high wages will on average be more favorable. Therefore, we consider the estimates obtained from our pooled cross-sections as an upper bound for the size of agglomeration economies, while we consider the estimates from our fixed effects estimates as a lower bound.

5 Results

5.1 Pooled Cross Sections

This section presents the estimation results of the regression models for workers with three different levels of education that have been described in Sect. 4.1, both using ordinary least squares (OLS so not controlling for potential endogeneity issues) and using instrumental variables (IV), as was discussed in Sect. 4.2. Results are presented in Table 2.

Even though the parameters estimated for individual worker characteristics are similar to what is generally found in the literature (see, for example, Groot et al. 2014), the results vary substantially across different levels of education. Low-educated females earn 18% less compared to their male colleagues with similar characteristics, while this figure is only 10% for females with a college degree. Also, the effect of age and thereby experience is much higher for high-educated workers: they thus follow a steeper career path compared to low-educated, as is shown in Fig. 3.

In contrast to the difference in the effect of age on wages that differ by level of education, the difference between part-time working and full-time employees is

Table 2 Regression results based on pooled cross-sections 2000–2010, OLS and IV

Dep.: log hourly wage	OLS (ordinary least squares)				IV (instrumental variables)			
	Low	Medium	College	University	Low	Medium	College	University
#Observations	3,755,289	3,225,783	2,877,818	1,641,720	3,736,950	3,211,272	2,868,281	1,639,025
#Employees	736,962	585,415	508,998	291,412	733,363	582,782	507,311	290,934
Females	−0.180 (37.9)	−0.173 (32.4)	−0.115 (29.1)	−0.100 (22.9)	−0.180 (37.7)	−0.173 (32.6)	−0.115 (28.7)	−0.099 (22.5)
Age	0.037 (67.3)	0.070 (66.8)	0.104 (101.8)	0.135 (73.6)	0.037 (66.5)	0.070 (66.5)	0.104 (100.6)	0.135 (74.0)
Age^2	−0.0004 (52.9)	−0.0007 (61.6)	−0.0011 (84.2)	−0.0014 (56.9)	−0.0004 (52.2)	−0.0007 61.4)	−0.0011 (83.4)	−0.0014 (57.5)
Part-time	−0.035 (15.1)	−0.084 (21.0)	−0.105 (34.9)	−0.132 (22.7)	−0.035 (15.2)	−0.083 (21.5)	−0.105 (34.9)	−0.132 (22.6)
Foreign (OECD)	−0.023 (10.4)	−0.030 (10.6)	−0.064 (10.7)	−0.091 (19.2)	−0.023 (10.3)	−0.032 (11.1)	−0.065 (11.0)	−0.092 (19.9)
Foreign (non-OECD)	–0.082 (18.2)	−0.111 (32.1)	−0.171 (53.1)	−0.233 (22.2)	−0.082 (17.8)	−0.112 (31.1)	−0.171 (53.6)	−0.234 (22.3)
Log employment density	0.040 (8.8)	0.065 (9.8)	0.065 (7.9)	0.087 (6.4)	0.028 (1.4)	0.089 (3.4)	0.083 (2.8)	0.112 (3.6)
Log share highly educated	−0.057 (1.9)	0.066 (1.4)	0.011 (0.2)	−0.105 (1.3)	0.055 (0.4)	−0.076 (0.4)	−0.118 (0.5)	−0.283 (1.2)
Year, sector and education dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.38	0.44	0.48	0.46	0.38	0.44	0.48	0.46

t values (OLS estimates) and *z* values (IV estimates) are in parentheses

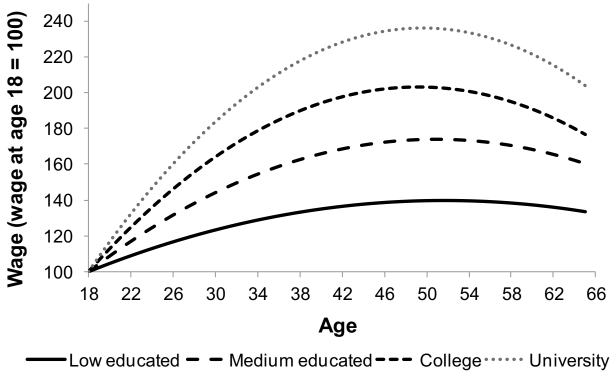

Fig. 3 Relation between age and expected wages by level of education

larger for high-educated workers. This can be explained by the fact that the type of jobs that can be performed with little loss of productivity per hour worked when working part-time are generally less skill intensive. For example, management positions are often difficult to perform part-time. The wage differential between foreign-born employees and natives is also substantially larger at higher levels of education, in particular for foreign employees that were born in non-OECD countries. This could be explained by the fact that high-skilled jobs require more communication and coordination, while differences in culture and language tend to work as a barrier when information needs to be exchanged between employees (cf. Groot 2013).

The results presented for regional level variables show that taking endogeneity issues into account has a large effect on the outcomes, as there are large differences between the parameters estimated for economic density and the share of high-educated employees estimated using OLS and those estimated with IV. The OLS estimates are thus substantially biased. When applying instrumental variables, the estimated agglomeration externalities are 2.8% for low-educated employees, 8.9% for medium skilled, 8.3% for college graduates and 11.2% for university graduates. Similar to the stylized facts presented in Sect. 3, we find that agglomeration externalities are relatively small for low-educated employees compared to medium and high-educated employees, while at the same time there is almost no difference in the importance of agglomeration for medium and high-skilled employees.

The average agglomeration elasticity (weighted by number of employees) across all levels of education is 7.1%. This is substantially higher as the estimates of, for example, Groot et al. (2014) who also use Dutch micro-data, but do not distinguish between workers with different levels of education. The average size of agglomeration externalities also seems to be at the upper end of the interval of 3–8% found by Melo et al. (2009) in their meta-analysis. The fact that we find the elasticity for lower educated employees to be below the lower boundary of Melo et al. (2009), while the estimates for medium and high-educated employees are above the upper boundary once more stresses the strong interaction between education and the importance of agglomeration economies.

An exogenous increase in the local supply of high-skilled workers that increases their share in the local labor market has a slightly positive effect on the wages of lower educated workers (when looking at the IV estimates), but a negative effect for both medium and high-educated employees. However, none of the estimated effects is statistically significant. Moretti (2004) also found higher social returns to education for lower skilled workers than for higher skilled workers, but he found a positive and statistically significant effect for high-educated workers as well. Our estimates thus show little evidence for the presence of positive knowledge spillovers in the Netherlands. Overall, the variables included in the regression model explain slightly less than half of the total variation in wages.

In the IV-specifications for low-skilled workers and college graduates in Table 2, the Kleibergen–Paap underidentification LM and Wald tests fail to reject their null hypotheses at a significance level of 0.10 (albeit barely), indicating some underidentification. In our IV-specifications with worker-fixed effects, instrument performance is much better with all underidentification tests rejected at $p=0.01$ or $p=0.001$.

5.2 Worker Fixed Effects

This section will compare the estimates presented in the previous section to similar estimates that include fixed effects on the level of individual employees to control for worker heterogeneity, as well as instruments for agglomeration and the share of highly educated employees. Because (for reasons discussed in Sect. 4.3) only employees whose work municipality changed over time are included in these regressions, the sample size is substantially reduced. The number of low-educated employees in our sample is reduced by 62%, that of medium educated by 54%, and that of college and university graduates by 50%, reflecting the fact that high-educated employees are generally more spatially mobile.

As Table 3 shows, the estimated agglomeration externalities are substantially lower when estimated while including fixed effects. The estimated elasticities are now estimated to be 2.7% for lower educated and 4.2% for all other levels of education. Even though the estimated agglomeration elasticities are smaller, the order and relative size has remained remarkably comparable to the IV estimates estimated on pooled cross-sections that were presented in Table 2. Again, we find that there is almost no difference in the relationship between agglomeration and wages for medium and high-educated employees, but a substantially lower effect of agglomeration on the wages of low-educated individuals. The magnitude of the estimated results is now much more comparable to the agglomeration externalities estimated in the previous literature for studies using micro data.

The relationship between the regional share of high-educated employees in the regional labor market and wages is now negative and statistically significant for workers with all levels of education, but this negative effect is much stronger for medium-educated workers as well as for workers with a college or university degree. The findings thus remain to be inconsistent with the literature predicting that there might be substantial social returns to education as measured by the local level of wages. Even though there exists a very strong correlation between individual level

Table 3 Regression results fixed effects estimation

Dep.: log hourly wage	OLS (ordinary least squares)				IV (instrumental variables)			
	Low	Medium	College	University	Low	Medium	College	University
#Observations	1,843,122	1,877,793	1,823,154	1,039,343	1,843,122	1,877,793	1,823,154	1,039,343
#Employees	278,008	268,740	256,006	146,028	278,008	268,740	256,006	146,028
Age	0.046 (79.5)	0.083 (71.4)	0.132 (99.0)	0.177 (64.5)	0.046 (63.7)	0.085 (56.9)	0.134 (82.0)	0.179 (59.5)
Age^2	−0.0004 (65.6)	−0.0008 (61.9)	−0.0013 (74.8)	−0.0017 (55.5)	−0.0004 (65.6)	−0.0007 (61.8)	−0.0013 (74.9)	−0.0017 (55.5)
Part-time	0.032 (30.0)	0.004 (2.6)	−0.017 (21.5)	−0.016 (10.6)	0.032 (30.1)	0.004 (2.6)	−0.017 (21.7)	−0.016 (10.6)
Log employment density	0.020 (12.8)	0.023 (15.2)	0.026 (12.7)	0.027 (9.7)	0.027 (5.4)	0.042 (5.6)	0.042 (5.2)	0.042 (4.0)
Log share highly educated	−0.050 (4.7)	−0.041 (3.4)	−0.042 (3.2)	−0.051 (2.6)	−0.094 (2.7)	−0.178 (3.6)	−0.157 (2.9)	−0.159 (2.1)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2 (within groups)	0.12	0.29	0.51	0.52	0.12	0.29	0.51	0.51

z values are in parentheses

wages and individual level of education, and between the average wage level in a region (even after controlling for observed worker heterogeneity using fairly detailed data) and the average level of education in a region, it is not likely that high average wages in a region are *caused* by a high share of the highly educated. A likely explanation for the difference with our findings is that the social returns estimated in some of the previous studies might have picked up some of the general relationship between education, agglomeration and wages.

Selection effects are a likely cause for the high share of high-educated workers in agglomerations. Because agglomeration externalities are larger for the highly-educated, this makes areas with a high economic density a particularly attractive place to work for them. Consistent with the laws of supply and demand, a high presence of other high-educated workers may in fact offset some of the advantages of working in a large agglomeration. We also find evidence for replacement effects of low-skilled by a higher supply of high-skilled workers.

6 Conclusion

The evidence found in this study has revealed substantial differences in the importance of agglomeration for employees with different levels of education. While a higher economic density of local labor markets brings only moderate advantages to low-educated workers, we find that both medium educated employees and college and university graduates earn substantially higher wages when they are employed in local labor markets with a higher employment density, even after controlling for observed and unobserved worker heterogeneity. Another important finding of the present study is that even though an exogenous increase in the share of high-educated workers results in a higher average wage in the region (because high-educated employees earn higher wages), it is negatively related to the wages of other employees in the region.

It is important to note that our findings by no means imply that the presence of universities and colleges has a negative impact on the wages in a region, as these institutions are one of driving forces behind agglomeration economies and the local knowledge infrastructure, thereby contributing to local productivity. For example, the constant supply of high-skilled employees is likely to make a region more attractive to firms that depend on the availability of high-skilled labor. It does, however, imply that at a given economic density and composition of the local labor market (with respect to firms and employees), employees will earn higher wages in regions without institutions in higher education.

Appendix 1: Estimating Distance Decay Functions to Define Local Labor Markets

The size and shape of spatial units that is used to estimate regional patterns in economic outcomes matters a great deal for the research findings (Briant et al. 2010; Combes and Gobillon 2015). The work of Groot et al. (2014), who estimate agglomeration externalities on both the level of Dutch municipalities and on the level of NUTS-3 regions, provides a good illustration of this phenomenon. On the NUTS-3 level, they find that doubling the employment density is associated with 4.8% higher regional wages (controlling for worker heterogeneity), while applying the exact same methodology on the level of municipalities yields an estimated agglomeration elasticity of only 2.1%. Briant et al. (2010) find that the optimal choice for a regional classification depends on the spatial scope of the phenomenon under investigation, whereby the level of spatial disaggregation should match the level at which the forces under examination are expected to operate. In the case of agglomeration forces, this is the level of the local labor market.

A further problem is that data availability often limits the options that researchers have when choosing the appropriate regional classification. Even when the level of detail is sufficient, availabilities are often restricted to administrative units which may deviate substantially from what can be considered a regional labor market. Besides the fact that the *average* size of such administrative areas may not be appropriate, there is also a substantial heterogeneity in the (spatial) size of regional units, particularly on the level of municipalities. Another problem arises from taking a discrete approach to defining a regional classification: if two individuals are located just a few meters apart but on different sides of the regional border they are considered to be in different regions (or in our case local labor markets), while in reality there is no real difference in location.

In this paper, we have therefore chosen to consider the relevant local labor market for an economic actor at a certain location as a continuum. The farther away from the core of each individual actors local labor market, the less an area is considered to be part of the relevant local labor market. Following Thompson (1965) and Horan and Tolbert (1984), we conceptually define local labor markets as the area around an economic core where labor market transactions generally take place, which is bounded by the radius within which most of the commuting towards the core takes place. A straightforward operational definition that follows from this theoretical definition is to consider the extent to which an area at a given radius from a location where economic activities take place is part of the relevant local labor market (in other words, the spatial weight of the area at that radius) to be equal to the cumulative distribution function of the fraction of commutes that take place within that radius or further. Thus, as 100% (10%) of commutes takes place within a radius of 0 km (50 km) or more, we apply a spatial weight of '1' ('0.1').

To formalize this relationship we have estimated a distance decay function. After experimenting with different functional specifications with up to three parameters, we found the following functional form to match the cumulative distribution function of observed commuting patterns almost exactly,

$$\begin{aligned} &w(r) = \alpha + \beta \cdot \ln(r) \quad \text{for} \quad r > 0 \quad \& \quad w > 0, \\ &w(r) = 0 \quad \text{for} \quad r = 0, \quad w(r) = 1 \quad \text{for} \quad w > 1, \end{aligned} \tag{3}$$

whereby w is the spatial weight of an area at a radius of distance r from the economic core. In the micro data that is available for this paper, we have both the residence and work municipality available for almost all Dutch employees (see Sect. 3

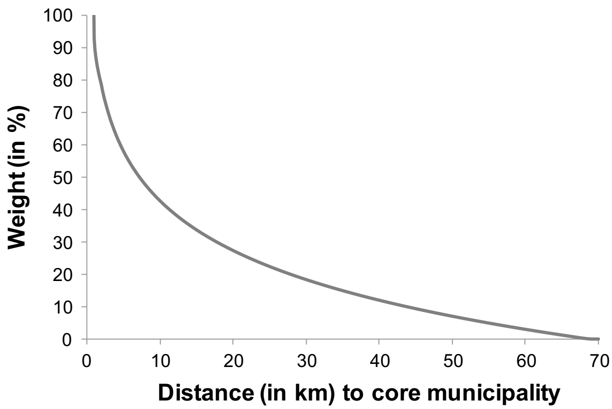

Fig. 4 Estimated distance decay function

for a discussion of our data), as well as the *x* and *y* coordinates of the center of each municipality. Using OLS to estimate the above equation resulted in parameter estimates of $\alpha = 0.9385$ and $\beta = -0.2219$. The relationship between the spatial weight and distance to the core municipality of the average Dutch local labor market is presented in Fig. 4. Even though 50% commutes less than 7 km, around 10% of all commuters live more than 40 km from their jobs. At a distance of 68 km the estimated distance decay function crosses the horizontal axis. Even though there is a small percentage of commutes within our data that takes place at distances up to 200 km, the fact that these commutes account for less than 1% of total commutes supports the view that the estimated cutoff point is appropriate.

Even though we could theoretically use actual the cumulative distribution function for each individual municipality as distance decay function, such a measure would be highly endogenous given our purpose of estimating agglomeration externalities. More productive regions characterized by high wages attract commuters from a very wide area compared to less productive and rural areas. Not in the least because of the increased demand for infrastructure that follows from these large commuting flows, there is generally more infrastructure connecting the large economic centers which results in better accessibility, attracting even more commuters (for this reason, estimating distance decay functions based on commuting time rather than distance is also problematic). If the distance decay function would be based on actual commuting towards a given municipality, the size of the spatial units would affect the size of agglomeration externalities, which is—given the findings of Briant et al. (2010) very likely to result in biased estimates. Therefore, we use the same distance decay parameters for all regions in our sample.

Figure 5 shows spatial weights for the local labor markets around three—for the purpose of illustration arbitrary chosen—municipalities: Amsterdam, Groningen and Maastricht. The weights quickly decline to 30–40% and decline more gradually from there onwards.

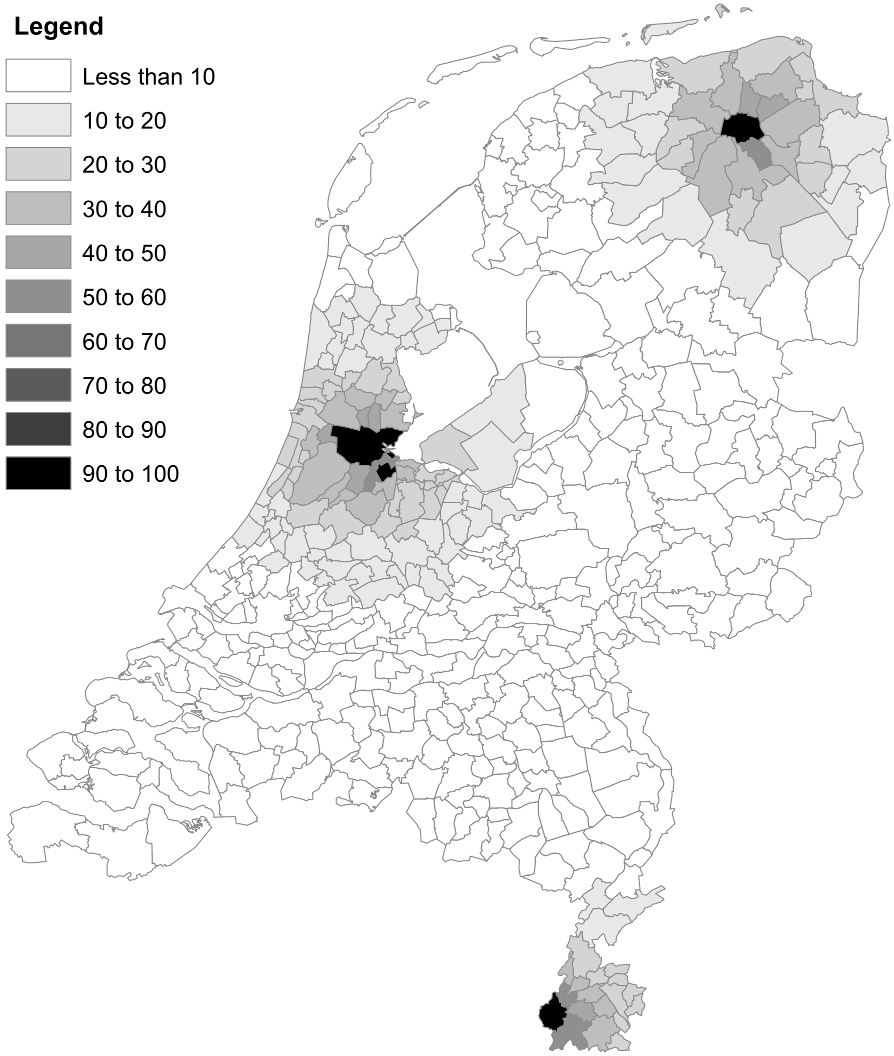

Fig. 5 Weights (in %) of municipalities around three local labor markets

References

Acemoglu, D., & Angrist, J. (2000). How large are human-capital externalities? Evidence from compulsory schooling laws. *NBER Macroeconomics Annual*, *15*, 9–59.

Beaudry, C., & Schiffauerova, A. (2009). Who's right, Marshall or Jacobs? The localization versus urbanization debate. *Research Policy, 38,* 318–337.

Briant, A., Combes, P. P., & Lafourcade, M. (2010). Dots to boxes: Do the size and shape of spatial units jeopardize economic geography estimations? *Journal of Urban Economics, 67,* 287–302.

Canton, E. (2009). Human capital externalities and proximity: Evidence from repeated cross-sectional data. *De Economist, 157,* 79–105.

Ciccone, A. (2000). Agglomeration effects in Europe. *European Economic Review, 46,* 213–227.

Ciccone, A., & Hall, R. E. (1996). Productivity and the density of economic activity. *American Economic Review, 86,* 54–70.
Combes, P. P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: sorting matters. *Journal of Urban Economics, 63,* 723–742.
Combes, P. P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2010). Estimating agglomeration effects with history, geology, and worker fixed-effects. In E. L. Glaeser (Ed.), *Agglomeration economics* (pp. 15–66). Chicago, IL: Chicago University Press.
Combes, P. P., & Gobillon, L. (2015). The empirics of agglomeration economies. *Handbook of Regional and Urban Economics, 5,* 247–348.
De Groot, H. L. F., Poot, J., & Smit, M. J. (2016). Which agglomeration externalities matter most and why? *Journal of Economic Surveys, 30,* 756–782.
De la Roca, J., & Puga, D. (2017). Learning by working in big cities. *The Review of Economic Studies*, *84*, 106–142.
Duranton, G. (2010). *The economics of clusters: Lessons from the French experience*. Oxford: Oxford University Press.
Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In V. Henderson & J. Thisse (Eds.), *Handbook of regional and urban economics* (pp. 2063–2117). Amsterdam: North-Holland.
Fujita, M. M. (1989). *Urban economic theory*. Cambridge: Cambridge University Press.
Fujita, M. M., Krugman, P., & Venables, A. J. (1999). *The spatial economy: Cities, regions, and international trade*. Cambridge, MA: MIT Press.
Fujita, M. M., & Thisse, J. (2002). *Economics of agglomeration-cities, industrial location and regional growth*. Cambridge: Cambridge University Press.
Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. *Journal of Political Economy, 100,* 1126–1152.
Glaeser, E. L., & Maré, D. C. (2001). Cities and skills. *Journal of Labor Economics, 19,* 316–342.
Graham, D. J., Melo, P. C., Jiwattanakulpaisarn, P., & Noland, R. B. (2010). Testing for bi-directional causality between productivity and agglomeration economies. *Journal of Regional Science, 50,* 935–951.
Groot, S. P. T. (2013). *Agglomeration, globalization and regional labor markets: Micro-evidence for the Netherlands, Tinbergen Institute Research Series, 553*. Amsterdam: Vrije Universiteit.
Groot, S. P. T., de Groot, H. L. F., & Smit, M. J. (2014). Regional wage differences in the Netherlands: Evidence, trends and explanations. *Journal of Regional Science, 54,* 503–523.
Groot, S. P. T., & Weterings, A. (2013). Internationalisation and firm productivity: Firm and regional level effects. In *CBS*, *internationalisation monitor 2013*, The Hague.
Harvey, D. (1981). The urban process under capitalism: A framework for analysis. In M. J. Dear & A. J. Scott (Eds.), *Urbanization and urban planning in capitalist society* (pp. 91–122). London: Taylor and Francis.
Henderson, J. V. (1988). *Urban development, theory, fact, and illusion*. Oxford: Oxford University Press.
Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics, 53,* 1–28.
Horan, P. M., & Tolbert, C. M. (1984). *The organization of work in rural and urban labor markets*. Boulder: Westview Press.
Jacobs, J. (1969). *The economy of cities*. New York: Random House.
Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *Quarterly Journal of Economics, 108,* 577–598.
Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics, 22,* 3–42.
Marshall, A. (1890). *Principles of economics*. London: Macmillan.
Melo, P. C., Graham, D. J., & Noland, R. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics, 39,* 332–342.
Mincer, J. (1974). *Schooling, experience, and earnings*. New York: NBER.
Moretti, E. (2004). Estimating the external return to education: Evidence from longitudinal and repeated cross-sectional data. *Journal of Econometrics, 121,* 175–212.
Moulton, B. R. (1990). An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics, 72,* 334–338.
Porter, M. E. (1990). *The competitive advantage of nations*. Houndmills: Macmillan.
Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science, 50,* 203–219.

Rauch, J. E. (1993). Productivity gains from geographic concentration of human capital: Evidence from the cities. *Journal of Urban Economics, 34,* 380–400.

Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In J. V. Henderson & J. Thisse (Eds.), *Handbook of urban and regional economics* (pp. 2119–2171). Amsterdam: Elsevier.

Schulz, T. P. (1988). Education investment and returns. In H. Chenery & T. N. Srinivasan (Eds.), *Handbook of development economics* (pp. 543–630). Amsterdam: North-Holland.

Smith, A. (1776). *The wealth of nations*. Lawrence: Digireads.com Publishing.

Thompson, W. R. (1965). *A preface to urban economics*. Baltimore: John Hopkins University Press.

Van Bergeijk, P. A. G., Fortanier, F., Garretsen, J. H., de Groot, H. L. F., & Moons, S. J. V. (2011). Productivity and internationalization: A micro-data approach. *De Economist, 159,* 381–388.

Verstraten, P., Verweij, G., & Zwaneveld, P. J. (2019). Complexities in the spatial scope of agglomeration externalities. *Journal of Regional Science, 59,* 29–55.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.